MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

K. KNOPP
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT
BERLIN

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE L. FEJÉR A. E. INGHAM
H. KNESER W. MAGNUS O. PERRON
W. SÜSS H. WEYL†

63. BAND, 5. (SCHLUSS-)HEFT

(ABGESCHLOSSEN AM 24. JANUAR 1956)

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1956

Math. Z.

Die **MATHEMATISCHE ZEITSCHRIFT**

wurde im Jahre 1918 von *L. Lichtenstein* unter der Mitwirkung von *K. Knopp*, *E. Schmidt* und *I. Schur* gegründet und herausgegeben. Nach dem Tode Lichtensteins übernahm *K. Knopp* 1933 die Herausgabe. Die Schriftleitung ergänzte sich 1933 durch *E. Kamke* und *F. K. Schmidt*, 1936 durch *R. Nevanlinna* und 1950 durch *H. Wielandt*, der 1952 die Herausgabe übernahm.

Die Zeitschrift erscheint nach Maßgabe des eingehenden Materials. Der Preis des Bandes beträgt DM 96.—.

Die Mathematische Zeitschrift ist durch jede Buchhandlung zu beziehen.

Die Mathematische Zeitschrift dient der Pflege der reinen Mathematik, doch werden auch Beiträge aus den Gebieten der theoretischen Physik und Astronomie Aufnahme finden, soweit sie mathematisch von Interesse sind. Besprechungen, Aufgaben u. dgl. werden nicht zugelassen.

Es wird ausdrücklich darauf aufmerksam gemacht, daß mit der Annahme des Manuskriptes und seiner Veröffentlichung durch den Verlag das ausschließliche Verlagsrecht für alle Sprachen und Länder an den Verlag übergeht, und zwar bis zum 31. Dezember desjenigen Kalenderjahres, das auf das Jahr des Erscheinens folgt. Hieraus ergibt sich, daß grundsätzlich nur Arbeiten angenommen werden können, die vorher weder im Inland noch im Ausland veröffentlicht worden sind, und die auch nachträglich nicht anderweitig zu veröffentlichen der Autor sich verpflichtet.

Die Mitarbeiter erhalten von ihren Arbeiten 75 Sonderdrucke unentgeltlich.

Manuskriptsendungen sind zu richten an die Mitglieder der unterzeichneten Redaktion. Die Herren Mitarbeiter werden im Interesse einer raschen Drucklegung gebeten, die Arbeiten in gut lesbarer Niederschrift einzureichen. Für den Text ist, wenn irgendmöglich, Maschinenschrift, für die mathematischen Formeln jedoch nur Handschrift zu verwenden. Etwaige Abbildungen sind auf einem besonderen Blatt zu zeichnen. Die in den Formeln etwa vorkommenden griechischen oder deutschen Buchstaben (Fraktur) sowie andere der Gefahr von Verwechselungen ausgesetzte Zeichen, z.B. die Landauschen Symbole o und O, sind stets besonders (etwa mit verschiedenfarbigen Stiften) zu kennzeichnen. Die Fußnoten sind fortlaufend zu numerieren, bei Zitaten Erscheinungsjahr und Seitenzahlen anzugeben.

Wenn zu einer bereits gedruckten Arbeit eine Berichtigung nötig wird, wollen die Herren Verfasser dem Herausgeber hiervon unverzüglich Mitteilung machen. Die Berichtigungen sollen dann am Schlusse eines Bandes gesammelt aufgenommen werden.

REDAKTION DER MATHEMATISCHEN ZEITSCHRIFT:

E. Kamke, Tübingen, Frischlinstraße 27,
K. Knopp, Tübingen, Am Apfelberg 6,
R. Nevanlinna, Helsinki (Finnland), Mäntytie 5B,
E. Schmidt, Berlin-Steglitz, Sedanstraße 8,
F. K. Schmidt, Heidelberg, Mathematisches Institut,
H. Wielandt, Tübingen, Ob dem Viehweidle 21b.

Springer-Verlag OHG — Berlin W 35 Reichpietschufer 20. — Heidelberg, Neuenheimer Landstraße 24. — Göttingen, Weender Straße 60.

---

63. Band **Inhalt** 5. Heft

Seite

Rademacher, H., Zur Theorie der Dedekindschen Summen 445
Stöhr, A., Zur Funktionentheorie im Raum der symmetrischen Matrizen 464
Wielandt, H., Primitive Permutationsgruppen vom Grad $2p$ 478
Rado, R., Minimal Points of Convex Sets in Sequence Spaces 486
Fenchel, W., Über konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten. 496
Kochendörffer, R., Über den Multiplikator einer Gruppe 507
Geiringer, H., Grenzlinien der Hodographentransformation 514
Mordell, L. J., The Minimum of an Inhomogeneous Quadratic Polynomial in n Variables 525
Brauer, A., On the Schnirelmann density of the sum of two sequences 529

RADEMACHER, H.
Math. Zeitschr. Bd. 63, S. 445—463 (1956)

Zur Theorie der Dedekindschen Summen

I. SCHUR zum Gedächtnis

Von

HANS RADEMACHER[1])

1. Die Dedekindschen Summen $s(h, k)$, die in der Theorie der Dedekindschen Funktion $\log \eta(\tau)$ auftreten, sind definiert als

$$s(h,k) = \sum_{\mu \bmod k} \left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) \qquad (h,k)=1, \quad k>0, \tag{1}$$

wo für reelles x

$$((x)) = \begin{cases} 0 \text{ für ganzes } x \\ x - [x] - \frac{1}{2} \text{ sonst} \end{cases} \tag{2}$$

gesetzt ist. Ihre bemerkenswerteste Eigenschaft wird durch die Reziprozitätsformel

$$s(h,k) + s(k,h) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h}{k} + \frac{k}{h} + \frac{1}{hk}\right) \tag{3}$$

ausgedrückt, die einerseits aus dem Verhalten von $\log \eta(\tau)$ unter Modulsubstitutionen abzulesen ist, für die aber auch eine Anzahl direkter Beweise bekannt sind. Man kann (2) mit Hilfe eines Euklidischen Algorithmus, da $s(h_1, k) = s(h_2, k)$, für $h_1 \equiv h_2 \pmod k$ ist, zweckmäßigerweise zur numerischen Berechnung von $s(h, k)$ für größere Werte von k, wo die Summendefinition unbequem wäre, benutzen.

Es treten in (2) und in einigen Verallgemeinerungen, die in verschiedener Richtung von MORDELL [*1*] und von mir [*2*] gegeben worden sind, jedoch mehrere $s(h, k)$ zugleich auf. Im folgenden will ich nun Eigenschaften studieren, die der einzelnen Summe $s(h, k)$ zukommen. Abschnitt I soll sich mit den Werten von $s(h, k)$ beschäftigen, Abschnitt II von einer Anwendung der Dedekindschen Summen auf das Problem der Transformationsklassen der Modulgruppe handeln, und Abschnitt III zieht quadratische Formen heran.

I

2. Hilfssatz 1. *Wenn h' durch*

$$h h' \equiv 1 \pmod k \tag{4}$$

bestimmt wird, ist

$$s(h', k) = s(h, k). \tag{4a}$$

Ferner ist

$$s(-h, k) = -s(h, k). \tag{4b}$$

[1]) John Simon Guggenheim Memorial Fellow 1954/55.

Beweis. Die zweite Aussage folgt sofort aus der Tatsache, daß $((x))$ eine ungerade Funktion von x ist.

Wenn ferner μ ein volles Restsystem modulo k durchläuft, so tut dies auch $h'\mu$. Wir haben also

$$s(h,k)=\sum_{\mu \bmod k}\left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right)=\sum_{\mu \bmod k}\left(\left(\frac{h'\mu}{k}\right)\right)\left(\left(\frac{hh'\mu}{k}\right)\right)=\sum_{\mu \bmod k}\left(\left(\frac{h'\mu}{k}\right)\right)\left(\left(\frac{\mu}{k}\right)\right)=s(h',k).$$

3. Da $((x))$ eine ungerade Funktion der Periode 1 ist, folgt sofort

$$\sum_{\lambda \bmod k}\left(\left(\frac{\lambda}{k}\right)\right)=0$$

und somit

$$s(h,k)=\sum_{\mu=1}^{k-1}\left(\frac{\mu}{k}-\frac{1}{2}\right)\left(\left(\frac{h\mu}{k}\right)\right)=\sum_{\mu=1}^{k-1}\frac{\mu}{k}\left(\left(\frac{h\mu}{k}\right)\right).$$

Explizit heißt dies

$$\text{(5)}\qquad \begin{cases} s(h,k)=\displaystyle\sum_{\mu=1}^{k-1}\frac{\mu}{k}\left(\frac{h\mu}{k}-\left[\frac{h\mu}{k}\right]-\frac{1}{2}\right) \\ \quad=\dfrac{h}{k^2}\displaystyle\sum_{1}^{k-1}\mu^2-\frac{1}{k}\sum_{1}^{k-1}\mu\left[\frac{h\mu}{k}\right]-\frac{1}{2k}\sum_{1}^{k-1}\mu \\ \quad=\dfrac{h(k-1)(2k-1)}{6k}-\dfrac{1}{k}g-\dfrac{k-1}{4}, \end{cases}$$

wo g eine ganze Zahl ist. Man sieht daher, daß der Nenner von $s(h,k)$ höchstens $6k$ sein kann. Da ferner in (5) der Nenner des einzelnen Summanden höchstens k^2 für gerades k und $2k^2$ für ungerades k beträgt, so ist der Nenner von $s(h,k)$ höchstens

für k gerade: $\qquad (6k,k^2)=k(6,k)=2k\cdot(3,k),$

für k ungerade: $\qquad (6k,2k^2)=2k(3,k).$

Damit haben wir erhalten den

Hilfssatz 2. *Der Nenner von $s(h,k)$ ist höchstens $2k\cdot(3,k)$.*

Dieser Wert des Nenners wird erreicht, z.B. ist $s(1,3)=\frac{1}{18}$. Der Nenner kann aber auch ein echter Teiler des angegebenen Wertes sein: $s(3,25)=\frac{2}{5}$.

Multiplikation von (3) mit $12hk$ ergibt

$$\text{(6)}\qquad 12hk\,s(h,k)+12hk\,s(k,h)=-3hk+h^2+k^2+1.$$

Im Hinblick auf Hilfssatz 2 lesen wir hieraus leicht ab den

Hilfssatz 3. *Es sei $\Theta=(3,k)$. Dann ist*

$$\text{(7)}\qquad 12hk\,s(h,k)\equiv h^2+1 \pmod{\Theta k}.$$

4. Wir beweisen nun den

Satz 1. *Der einzige ganzzahlige Wert, den $12s(h,k)$ annimmt, ist Null. Dies tritt dann und nur dann ein, wenn*

$$\text{(8)}\qquad h^2+1\equiv 0 \pmod{k}.$$

Beweis. Es sei zunächst (8) erfüllt. Dann ist, in der Bezeichnung Hilfssatz 1

$$h' = -h,$$

also

$$s(h,k) = s(-h,k) = -s(h,k),$$

woraus die zweite Hälfte des Satzes folgt.

Sei nun umgekehrt $12s(h,k)$ ganz. Dann ist nach (6) und Hilfssatz 2 die Kongruenz (8) erfüllt.

5. Satz 2. *Es ist* $s(h,k) < s(1,k)$ *für* $1 < h < k$.

Beweis. Wir haben nach (5)

$$k \cdot s(1,k) = \sum_{\mu=1}^{k-1} \mu\left(\frac{\mu}{k} - \frac{1}{2}\right),$$

$$k \cdot s(h,k) = \sum_{\mu=1}^{k-1} \mu\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right] - \frac{1}{2}\right).$$

Es genügt also zu zeigen, daß

$$\sum_{\mu=1}^{k-1} \mu^2 > \sum_{\mu=1}^{k-1} \mu\left(h\mu - k\left[\frac{h\mu}{k}\right]\right)$$

ist. Hier sind die Zahlen

$$a_\mu = h\mu - k\left[\frac{h\mu}{k}\right]$$

eine Permutation der Zahlen $1, 2, \ldots, k-1$. Für alle möglichen Permutationen $\{a_\mu\}$ der Zahlen 1 bis $k-1$ erreicht aber die Summe

$$\sum \mu a_\mu$$

ihr Maximum für die Anordnung $a_\mu = \mu$. In der Tat, wenn nicht durchgängig $\mu = a_\mu$ ist, dann gibt es in der Anordnung $a_1, a_2, \ldots, a_{k-1}$ ein größtes $\mu = \mu_0$, so daß $a_{\mu_0} \neq \mu_0$, während $a_\mu = \mu$ für $\mu > \mu_0$ (sofern ein solches μ vorrätig ist). Dann ist jedoch sogar $a_{\mu_0} < \mu_0$, denn die $a_\mu > \mu_0$ sind besetzt durch $a_\mu = \mu$. Es sei dann $\mu_0 = a_{\mu_1}$ mit $\mu_1 < \mu_0$. Wir vertauschen dann a_{μ_0} und a_{μ_1}, ersetzen also die beiden Summanden $\mu_1 a_{\mu_1} + \mu_0 a_{\mu_0}$ durch $\mu_1 a_{\mu_0} + \mu_0 a_{\mu_1}$, was eine Vergrößerung bedeutet:

$$\mu_1 a_{\mu_0} + \mu_0 a_{\mu_1} - (\mu_1 a_{\mu_1} + \mu_0 a_{\mu_0}) = (\mu_0 - \mu_1)(a_{\mu_1} - a_{\mu_0}) = (\mu_0 - \mu_1)(\mu_0 - a_{\mu_0}) > 0.$$

Satz 3. *Es ist* $s(h,k) > 0$ *für* $0 < h < \sqrt{k-1}$.

Beweis. Nach (3) und nach Satz 2 haben wir

$$s(h,k) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h^2+1}{hk} + \frac{k}{h}\right) - s(k,h)$$

$$\geqq -\frac{1}{4} + \frac{1}{12}\left(\frac{h^2+1}{hk} + \frac{k}{h}\right) - s(1,h).$$

Nun folgt aus (3), da $s(h, 1) = 0$ ist,

$$s(1, h) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h^2+1}{h} + \frac{1}{h}\right),$$

so daß wir erhalten

$$s(h, k) \geq \frac{(k-1)(k-h^2-1)}{12hk},$$

woraus die Behauptung des Satzes folgt.

Da in dieser Überlegung das Ungleichheitszeichen nur bei der Ersetzung von $s(k, h)$ durch $s(1, h)$ auftritt, so haben wir das

Korollar. $s(h, k) = \dfrac{(k-1)\left(\dfrac{k-1}{h} - h\right)}{12k}$ *für* $k \equiv 1 \pmod{h}$.

5. Satz 4. *Es seien* $\frac{h_1}{k_1} < \frac{h_2}{k_2}$ *zwei benachbarte Brüche in einer Farey-Reihe. Dann ist*

$$s(h_1, k_1) - s(h_2, k_2) = \frac{1}{4} - \frac{1 + k_1^2 + k_2^2}{12 k_1 k_2}. \tag{9}$$

Beweis. Die kennzeichnende Eigenschaft für benachbarte Farey-Brüche $\frac{h_1}{k_1} < \frac{h_2}{k_2}$ ist

$$\begin{vmatrix} h_1 & h_2 \\ k_1 & k_2 \end{vmatrix} = -1,$$

so daß also

$$h_1 k_2 \equiv -1 \pmod{k_1},$$
$$h_2 k_1 \equiv 1 \pmod{k_2}.$$

Dann ist aber nach Hilfssatz 1

$$s(k_2, k_1) = -s(h_1, k_1),$$
$$s(k_1, k_2) = s(h_2, k_2),$$

womit (9) aus (3) folgt, wenn dort h, k durch k_1, k_2 ersetzt werden.

Satz 5. *Über* h_1/k_1 *und* h_2/k_2 *mögen die Voraussetzungen von Satz 4 gelten. Dann ist*

$$s(h_1 + h_2, k_1 + k_2) = \frac{1}{2}\{s(h_1, k_1) + s(h_2, k_2)\} + \frac{(k_2 - k_1)(1 + k_1^2 + k_1 k_2 + k_2^2)}{24 k_1 k_2 (k_1 + k_2)}. \tag{10}$$

Beweis. Die Anwendung von (9) auf $\frac{h_1}{k_1} < \frac{h_1+h_2}{k_1+k_2} < \frac{h_2}{k_2}$ ergibt

$$s(h_1, k_1) - s(h_1 + h_2, k_1 + k_2) = \frac{1}{4} - \frac{1 + k_1^2 + (k_1 + k_2)^2}{12 k_1 (k_1 + k_2)},$$

$$s(h_1 + h_2, k_1 + k_2) - s(h_2, k_2) = \frac{1}{4} - \frac{1 + (k_1 + k_2)^2 + k_2^2}{12 (k_1 + k_2) k_2}.$$

Durch Subtraktion dieser beiden Gleichungen voneinander ergibt sich (10).

Die Formel (10) eignet sich besonders gut zur schrittweisen tabellarischen Berechnung von $s(h, k)$, ausgehend von $s(0, 1) = s(1, 1) = 0$.

6. Die Sätze 4 und 5 legen nahe, $s(h/k)$ statt $s(h, k)$ zu schreiben. Dies kann man unbedenklich tun, wenn man vorübergehend die Bedingung $(h, k) = 1$ fallen läßt. In der Tat hängt dann $s(h, k)$ nur von dem Verhältnis $h:k$ ab, wie man ersieht aus

Hilfssatz 4. *Für ganzes positives q ist*

$$s(qh, qk) = s(h, k).$$

Dies gilt nur für die Definition (1) von $s(h, k)$, nicht etwa für Formel (5). Dieser Hilfssatz ist Theorem 1 in [3] und dort bewiesen. [Alle Formeln setzen aber bisher $(h, k) = 1$ voraus.] Damit ist eine Funktion $s(\varrho)$ für alle rationalen $\varrho = h/k$ definiert. Über den Verlauf dieser Funktion x macht der folgende Satz eine Aussage.

Satz 6. *Die für rationale $\varrho = h/k$ definierte Funktion $s(\varrho)$ ist in jedem Intervall nach oben und unten unbeschränkt.*

Beweis. In (9) möge h_1/k_1 festgehalten werden und h_2/k_2 eine Folge von benachbarten Brüchen durchlaufen die von oben gegen h_1/k_1 gehen (z.B. durch festgesetze Mediantenbildung). Mit k_1 fest, $k_2 \to +\infty$ geht die rechte Seite von (9) gegen $-\infty$, also $s(h_2/k_2) \to +\infty$. Ebenso, wenn man h_2/k_2 festhält und h_1/k_1 gegen h_2/k_2 gehen läßt, folgt $s(h_1/k_1) \to -\infty$.

II

7. Es sei Γ die Gruppe aller homogenen Modulsubstitutionen

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 1.$$

Zwei Substitutionen M_1, M_2 heißen ähnlich, wenn ein $L \in \Gamma$ existiert, so daß

$$M_1 = L^{-1} M_2 L.$$

Die Gruppe Γ zerfällt durch die Ähnlichkeitsrelation in *Klassen*. Die Spur $a + d$ ist offenbar eine Invariante der Ähnlichkeitsklassen. In diesem Abschnitt behandeln wir eine weitere Invariante.

Satz 7. *Die Funktion*

$$(11) \quad \begin{cases} \Psi(M) = \Psi\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ \qquad = \begin{cases} \dfrac{b}{d} & \text{für } c = 0, \\ \dfrac{a+d}{c} - 12 \operatorname{sign} c \cdot s(a, |c|) - 3 \operatorname{sign}(c(a+d)) & \text{für } c \neq 0 \end{cases} \end{cases}$$

ist eine Klasseninvariante. Außerdem ist

$$(12) \qquad \Psi(M) = \Psi(-M), \quad \Psi(M^{-1}) = -\Psi(M),$$

und Ψ ist eine ganze Zahl.

Beweis. Die Aussagen (12) folgen aus der Definition (11), wenn man noch

$$M^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

beachtet. Ferner kann man aus (7) entnehmen

$$12|c|\, s(a,|c|) \equiv a + d \pmod{c},$$

also

$$\frac{a+d}{|c|} - 12\, s(a,|c|)$$

ganzzahlig, woraus die Ganzzahligkeit von Ψ in (11) folgt.

Für die Invarianz von $\Psi(M)$ brauchen wir nur die Gleichungen

$$\Psi(S^{-1}MS) = \Psi(M), \tag{13}$$

$$\Psi(T^{-1}MT) = \Psi(M), \tag{14}$$

mit

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

zu beweisen.

Mit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

haben wir

$$S^{-1}MS = \begin{pmatrix} a-c & a-c+b-d \\ c & c+d \end{pmatrix}, \tag{13a}$$

$$T^{-1}MT = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \tag{14a}$$

Dann ist (13) unmittelbar aus der Definition von Ψ abzulesen. Wir haben also nur noch (14) zu untersuchen. Zunächst bemerken wir, daß

$$12\, s(1,k) = -3 + k + \frac{2}{k}, \tag{15}$$

was aus dem Korollar zu Satz 3, oder auch direkt aus (3) folgt. Für den Beweis von (14) unterscheiden wir drei Fälle.

I. $c = 0$, dann ist $a = d = 1$. [$a = d = -1$ kann auf Grund von (12) beiseite gelassen werden.] Dann lautet (14), (14a):

$$\Psi\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \Psi\begin{pmatrix} 1 & 0 \\ -b & 1 \end{pmatrix} \tag{16}$$

oder, nach (11),

$$b = \begin{cases} 0 & \text{für } b = 0 \\ \dfrac{2}{-b} + 12 \operatorname{sign} b \cdot s(1,|b|) + 3 \operatorname{sign} b, \end{cases}$$

was in der Tat durch (15) für $k = |b|$ bestätigt wird.

II. $c \neq 0$, $b = 0$. Dann handelt es sich um

$$\Psi\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} = \Psi\begin{pmatrix} 1 & -c \\ 0 & 1 \end{pmatrix},$$

was aus (16) durch Vertauschung von c und $-b$ hervorgeht und somit unter I erledigt ist.

III. $b \neq 0$, $c \neq 0$. Hier ist also im Hinblick auf (11), (14), (14a) zu beweisen:

$$(17) \quad \begin{cases} 12 \operatorname{sign} c \cdot s(a, |c|) + 12 \operatorname{sign} b \cdot s(d, |b|) \\ = \dfrac{a+d}{c} + \dfrac{a+d}{b} - 3 \operatorname{sign}\big(c(a+d)\big) - 3 \operatorname{sign}\big(b(a+d)\big). \end{cases}$$

Wenn $d = 0$ ist, dann muß $b = -c = \pm 1$ sein, und (17) ist dann trivialerweise erfüllt.

Wenn $d \neq 0$, dann kann wegen (12) $d > 0$ angenommen werden. Dann ergibt (3)

$$12\, s(d, |c|) + 12\, s(|c|, d) = -3 + \frac{d}{|c|} + \frac{|c|}{d} + \frac{1}{d|c|},$$

$$12\, s(d, |b|) + 12\, s(|b|, d) = -3 + \frac{d}{|b|} + \frac{|b|}{d} + \frac{1}{d|b|}.$$

Die erste dieser Gleichungen multiplizieren wir mit $\operatorname{sign} c$, die zweite mit $\operatorname{sign} b$ und beachten, daß zufolge (4b)

$$s(h, k) = \operatorname{sign} h \cdot s(|h|, k)$$

ist, so daß

$$(18) \quad \begin{cases} 12 \operatorname{sign} c \cdot s(d, |c|) + 12\, s(c, d) = -3 \operatorname{sign} c + \dfrac{d}{c} + \dfrac{c}{d} + \dfrac{1}{dc}, \\ 12 \operatorname{sign} b \cdot s(d, |b|) + 12\, s(b, d) = -3 \operatorname{sign} b + \dfrac{d}{b} + \dfrac{b}{d} + \dfrac{1}{db} \end{cases}$$

hervorgeht. Nach Hilfssatz 1 ist

$$s(a, |c|) = s(d, |c|),$$

$$s(c, d) = -s(b, d).$$

Addition der Gln. (18) ergibt also

$$(19) \quad \begin{cases} 12 \operatorname{sign} c \cdot s(a, |c|) + 12 \operatorname{sign} b \cdot s(d, |b|) \\ = -3 \operatorname{sign} c - 3 \operatorname{sign} b + \dfrac{d}{c} + \dfrac{c}{d} + \dfrac{ad - bc}{dc} + \dfrac{d}{b} + \dfrac{b}{d} + \dfrac{ad - bc}{db}. \end{cases}$$

Ein Vergleich von (17) und (19) zeigt, daß (17) bewiesen sein wird, sobald die Richtigkeit von

$$(20) \qquad \operatorname{sign}\big(c(a+d)\big) + \operatorname{sign}\big(b(a+d)\big) = \operatorname{sign} c + \operatorname{sign} b$$

gezeigt ist. Nun ist (20) klar für $a+d>0$. Für $a+d\leq 0$ und $d>0$ (was wir annehmen durften) folgt $ad<0$, und somit $bc=ad-1<0$, was entgegengesetzte Vorzeichen von b und c bedeutet. Dann sind beide Seiten von (20) gleich Null. Somit ist (17) bewiesen und damit der Beweis von Satz 7 erbracht.

8. Unter Einführung der Matrix $U=ST$ mit

$$T^2=U^3=1 \tag{21}$$

(wo Faktoren ± 1 vernachlässigt, d.h. die inhomogenen Modulsubstitutionen gemeint sind) kann jede Modulsubstitution *eindeutig* in kürzester Form dargestellt werden[2]) als

$$M=U^{\varepsilon_0}TU^{\varepsilon_1}T\dots TU^{\varepsilon_{\nu+1}}, \tag{22}$$

wo

$$\varepsilon_j=\pm 1, \quad j=1,\dots,\nu$$
$$\varepsilon_0, \varepsilon_{\nu+1}=0 \quad \text{oder} \quad \pm 1.$$

Elemente M, die durch zyklische Vertauschung der Faktoren der Darstellung (22) auseinander hervorgehen, gehören zur selben Klasse. Wenn man das Produkt (22) zyklisch schließt, wird im allgemeinen eine Reduktion durch Anwendung von (21) eintreten. Dann gibt es also in einer Klasse immer eine Transformation von folgender Art

1. die elliptischen T, U, U^{-1}

oder

2. $TU^{\varepsilon_1}TU^{\varepsilon_2}\dots TU^{\varepsilon_\nu}, \quad \varepsilon=\pm 1.$

Hilfssatz 4. *Ist*

$$M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}=TU^{\varepsilon_1}TU^{\varepsilon_2}\dots TU^{\varepsilon_\nu}, \quad \varepsilon_j=\pm 1, \tag{23}$$

dann gilt[3])

$$a<0, \quad d<0, \quad b\geq 0, \quad c\geq 0. \tag{24}$$

Beweis. Diese Behauptungen sind richtig für $\nu=1$, also für

$$TU=TST=\begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix},$$
$$TU^{-1}=S^{-1}=\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Angenommen die Behauptungen (24) wären für M in der Form (23) gültig, dann haben wir zu zeigen, daß sie auch für

$$M\cdot TU^{\varepsilon_{\nu+1}}$$

[2]) Den Hinweis auf die Bedeutung der kürzesten Darstellung in diesem Zusammenhang verdanke ich Herrn REIDEMEISTER; vgl. [4], S. 42, 43.

[3]) Möglicherweise zu erreichen durch Multiplikation der Matrix mit -1.

zutreffen. Nun ist

$$M\,T\,U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -a+b & -b \\ -c+d & -d \end{pmatrix} = \begin{pmatrix} a-b & b \\ c-d & d \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

und

$$M\,T\,U^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -a & a-b \\ -c & c-d \end{pmatrix} = \begin{pmatrix} a & b-a \\ c & d-c \end{pmatrix} = \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix}.$$

Aus (24) folgt dann dieselbe Behauptung auch für a', b', c', d', und a'', b'', c'', d'', womit der Hilfssatz bewiesen ist.

9. Nunmehr kann man der Klasseninvariante Ψ eine neue Bedeutung geben. Es gilt nämlich der

Satz 8. *Es sei M in seiner Klasse in der zyklisch reduzierten Form angenommen. Dann gilt:*

$$(25)\qquad \begin{cases} \Psi(T) = \Psi\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = 0, \\ \Psi(U) = \Psi\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = -2, \quad \Psi(U^{-1}) = \Psi\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} = 2 \end{cases}$$

und im Falle (23)

$$(26)\qquad \Psi(M) = \sum_{j=1}^{\nu} \varepsilon_j.$$

Beweis. Die Formeln (25) liest man direkt aus der Definition (11) von Ψ ab. Die Gl. (26) wird durch die für den Beweis von Hilfssatz 4 angewandte Induktion bewiesen.

Zunächst ist

$$\Psi(T\,U) = \Psi\begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} = 1,$$

$$\Psi(T\,U^{-1}) = \Psi\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = -1,$$

wie aus (11) folgt. Also stimmt (26) für $\nu = 1$.

Es sei nun (26) für ein ν erfüllt, dann bleibt nur zu zeigen, daß

$$(27)\qquad \Psi(M\,T\,U^{\varepsilon_{\nu+1}}) = \Psi(M) + \varepsilon_{\nu+1}, \qquad \varepsilon_{\nu+1} = \pm 1$$

ist. Sei

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

dann ist

$$M\,T\,U = \begin{pmatrix} a-b & b \\ c-d & d \end{pmatrix},$$

also

$$\Psi(M) = \begin{cases} \frac{b}{d} & \text{für} \quad c = 0 \\ \frac{a+d}{c} - 12s(d,c) + 3, & c > 0, \end{cases} \tag{28a}$$

$$\Psi(M\,T\,U) = \frac{a-b+d}{c-d} - 12s(d, c-d) + 3, \tag{28b}$$

wo nach (24) $c-d=0$ nicht vorkommt, und wo, wieder nach (24), $a+d<0$, $a-b+d<0$ benutzt ist.

Sei zunächst $c=0$, also $a=d=-1$. Dann ist

$$\begin{cases} \Psi(M) = \frac{b}{d} = -b, & \Psi(M\,T\,U) = \frac{-1-b-1}{-d} + 3 \\ & \qquad = -b+1, \end{cases} \tag{29}$$

in Übereinstimmung mit (27).

Sei zweitens $c>0$, also nach (28a), (28b)

$$\Psi(M\,T\,U) - \Psi(M) = \frac{a-b+d}{c-d} - \frac{a+d}{c} - 12s(d, c-d) + 12s(d,c).$$

Nun ist aber nach (3)

$$-12s(d, c-d) = 12s(c,d) + 3 - \frac{d^2 + (c-d)^2 + 1}{d(c-d)}$$

und somit

$$\Psi(M\,T\,U) - \Psi(M) = 12s(c,d) + 12s(d,c) + \\ + 3 - \frac{d^2 + (c-d)^2 + 1}{d(c-d)} + \frac{a-b+d}{c-d} - \frac{a+d}{c}.$$

Eine nochmalige Anwendung von (3) ergibt dann

$$\Psi(M\,T\,U) - \Psi(M) = \frac{c^2 + d^2 + 1}{cd} - \frac{a+d}{c} + \frac{b-c+d}{d} = 1.$$

Diese Gleichung und (29) bestätigen (27) für den Fall $\varepsilon_{\nu+1} = +1$.

Ferner ist

$$\Psi(M\,T\,U^{-1}) = \Psi\begin{pmatrix} a & -a+b \\ c & -c+d \end{pmatrix} = \begin{cases} -\frac{a+b}{d} & c=0 \\ \frac{a-c+d}{c} - 12s(d,c) + 3, & c>0 \end{cases} \tag{30}$$

und somit für $c=0$, $a=d=-1$:

$$\Psi(M) = -b, \qquad \Psi(M\,T\,U^{-1}) = -1-b,$$

während sich für $c>0$ aus (28a) und (30)

$$\Psi(M\,T\,U^{-1}) - \Psi(M) = -\frac{c}{c} = -1$$

ergibt. Damit ist (27) auch für den Fall $\varepsilon_{\nu+1} = -1$ bewiesen und der Induktionsbeweis von Satz 8 zu Ende geführt.

10. Im allgemeinen kann man nur aussagen, daß $\Psi(M_1 M_2)$ und $\Psi(M_1) + \Psi(M_2)$ sich nur um Vielfache von 3 unterscheiden können. Jedoch gilt der

Satz 9. *Wenn M keine elliptische Substitution ist, dann ist für ganzes k*

$$\Psi(M^k) = k\,\Psi(M).$$

Beweis. Dies folgt zunächst für $k > 0$ daraus, daß in dem zyklischen Zusammenschluß von M_k k mal die gleichen Reduktionen auftreten, die nur einmal bei dem zyklischen Zusammenschluß von M auftreten. Also besteht das zyklisch reduzierte M_k aus k Wiederholungen des zyklisch reduzierten M, was den Satz für $k > 0$ beweist. Für $k < 0$ folgt der Satz dann aus der zweiten Gl. (12), und für $k = 0$ aus $\Psi(E) = 0$, wo E die Einheitsmatrix ist.

In etwas anderer Richtung liegt der

Satz 10. *Wenn für*

$$M_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}, \tag{31}$$

die Ungleichungen

$$\begin{cases} a_1 < 0,\ d_1 < 0, & c_1 > 0,\ b_1 > 0 \\ a_2 < 0,\ d_2 < 0, & c_2 > 0,\ d_2 > 0 \end{cases} \tag{32}$$

gelten, so ist

$$\Psi(M_1 M_2) = \Psi(M_1) + \Psi(M_2). \tag{33}$$

Beweis. Die Ungleichungen (32) sorgen dafür, daß in der Normaldarstellung (22) M_1 so endet und M_2 so anfängt, daß keine gegenseitigen Reduktionen in dem Produkt $M_1 M_2$ stattfinden.

Um dies einzusehen, ändern wir die Bedingungen zu (22) dahin ab, daß wir nur nicht-negative Potenzen zulassen, was dadurch geschieht, daß jeder etwa auftretende Exponent -1 durch 2 ersetzt wird. Dann werde die Definition $U = ST$ benutzt, so daß (22) umgeformt wird in

$$\begin{cases} M = (ST)^{r_0} T (ST)^{r_1} T \dots (ST)^{r_\nu} T (ST)^{r_{\nu+1}}, \\ \qquad r_0, r_{\nu+1} = 0, 1, 2, \quad r_j = 1, 2, \quad j = 1, \dots, \nu. \end{cases} \tag{34}$$

Dann und nur dann, wenn $r_0 = 0$, $r_{\nu+1} > 0$ wird M in (22) die spezielle Form (23) haben, d.h. mit T anfangen und mit einer Potenz von U enden. Wir wollen zeigen, daß aus

$$a < 0, \quad d < 0, \quad c > 0 \tag{35}$$

folgt $r_0 = 0$, $r_{\nu+1} > 0$.

In der Tat, (34) kann nach gehörigem Zusammenziehen auf Grund von $T^2 = 1$ geschrieben werden:

$$M = S^{q_0} T S^{q_1} \dots T S^{q_l} T S^{q_{l+1}}, \quad q_j \gtrless 1,\ j = 1, \dots l, \quad q_0 \geqq 0,\ q_{l+1} \geqq 0, \tag{36}$$

wo $q_0 > 0$ aus $\eta_0 > 0$, und $q_{l+1} > 0$ aus $\eta_{\nu+1} = 0$ folgen würde. Man sieht ferner leicht, daß wegen (34) von den nicht-verschwindenden q_j nur das erste und letzte gleich 1 sein können, während für die übrigen sogar $q_j \geqq 2$ gilt. Nach (36) kann man

$$M(\tau) = \frac{a\tau + b}{c\tau + d}$$

als einen Kettenbruch

$$\frac{a\tau+b}{c\tau+d} = q_0 - \cfrac{1}{q_1 - \ddots - \cfrac{1}{q_l - \cfrac{1}{q_{l+1} + \tau}}}$$

schreiben, woraus durch $\frac{1}{\tau} \to 0$

$$\frac{a}{c} = q_0 - \cfrac{1}{q_1 - \ddots - \cfrac{1}{q}}$$

folgt, so daß $q_0 > 0$

$$\frac{a}{c} \geqq 0$$

nach sich ziehen würde, gegen die Ungleichungen (35). Ebenso beweist man aus

$$M^{-1} = S^{-q_{l+1}} T S^{-q} \dots T S^{-q_1} T,$$

also aus

$$-\frac{d}{c} = -q_{l+1} - \cfrac{1}{-q_l - \ddots - \cfrac{1}{-q_1}},$$

daß $q_{l+1} > 0$ die Ungleichung

$$-\frac{d}{c} \leqq 0$$

zur Folge hätte, wieder gegen (35). Es ist also $q_0 = q_{l+1} = 0$ und folglich $\eta_0 = 0, \eta_{\nu+1} > 0$, und M hat die Form (23), wobei nur $\varepsilon_{\nu+1}$ statt ε_ν zu schreiben ist.

Daher haben infolge von (32) die Substitutionen M_1 und M_2 auch die Form (23). Dann aber tritt in dem Produkt $M_1 M_2$ keine Reduktion ein, da auf das letzte U^{ε_ν} von M, das erste T von M_2 folgt. Damit ist Satz 10 bewiesen.

Wenn man in (33) die Matrizen von (31) einträgt und die Definition (11) von Ψ aus Satz 7 heranzieht, so folgt nach einigen Rechnungen das

Korollar. *Unter den Bedingungen* (32) *ist*

$$s(d, |c|) + s(d_1, c_1) + s(d_2, c_2) = \frac{1}{4} - \frac{c_1^2 + c_2^2 + c^2}{12 c_1 c_2 |c|}$$

mit

$$c = c_1 a_2 + d_1 c_2, \quad d = c_1 b_2 + d_1 d_2.$$

Diese Formel könnte auch analytisch aus der Theorie der Funktion $\log \eta(\tau)$ gewonnen werden. Sie folgt übrigens auch aus der dreigliedrigen Formel (30) in [2].

III

11. Für die elliptischen Substitutionen, die mit T oder U oder U^{-1} äquivalent sind, also nur zu 3 Klassen gehören können, ist die Spur $a+d$ gleich 0 oder ± 1. Wenn man die Spur $a+d=m$ mit $|m| \geq 2$ vorgibt, wobei man $a+d$ negativ wählen kann, so kann man nach dem Hilfssatz 4 und den ihm vorausgehenden Erörterungen in derselben Klasse stets eine Substitution in der Form (23) finden, für die $a<0$, $d<0$ ist. Solche a, d gibt es aber bei festem $a+d=m$ nur endlich viele.

Ist $m \leq -3$, so ist $a \cdot d \neq 1$, also

$$bc = ad - 1 \neq 0,$$

und von den b, c (die beide nach Hilfssatz 4 positiv zu sein haben), gibt es daher wieder nur endlich viele. Also gibt es mit vorgegebener Spur $a+d=m \leq -3$ nur je endlich viele Klassen von Modulsubstitutionen.

Für $a+d=-2$, also $a=d=-1$, $bc=0$ gibt es die unendlich vielen Substitutionen

$$(TU)^c = \begin{pmatrix} -1 & 0 \\ c & -1 \end{pmatrix}, \qquad c>0,$$

$$(TU^{-1})^b = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix}, \qquad b>0,$$

die verschiedene Normalformen (23) zeigen, und also zu verschiedenen Klassen gehören. Dies sind die parabolischen Substitutionen. Es ist übrigens

$$\Psi\left((TU)^c\right) = c, \qquad \Psi\left((TU^{-1})^b\right) = -b.$$

12. Einen anderen Einblick in die Verteilung auf die Klassen bei gegebenem $m=a+d$ gewinnt man durch eine Zuordnung von quadratischen Formen zu den Modulsubstitutionen.

Wir gehen dazu wieder zu den homogenen Modulsubstitutionen über, unterscheiden also die unimodularen Matrizen M und $-M$. Dann werde der Modulsubstitution

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

die quadratische Form

(36) $$Q(x,y) = cx^2 + (d-a)xy - by^2$$

mit der Diskriminante

$$\Delta = (d-a)^2 + 4bc = (a+d)^2 - 4 = m^2 - 4$$

zugeordnet. Es ist klar, daß zu

$$-M^{-1} = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix}$$

dieselbe Form Q gehört wie zu M.

Umgekehrt, hat

(37) $$Q(x, y) = A x^2 + B x y + C y^2$$

die Diskriminante

(38) $$\Delta = B^2 - 4AC = m^2 - 4,$$

dann gehört sie zu

$$M = \begin{pmatrix} \frac{1}{2}(m - B) & -C \\ A & \frac{1}{2}(m + B) \end{pmatrix} \text{ und } -M^{-1} = \begin{pmatrix} -\frac{1}{2}(m + B) & -C \\ A & \frac{1}{2}(m - B) \end{pmatrix}.$$

Die Elemente in M sind ganz, da aus (38)

$$B \equiv m \pmod 2$$

folgt.

Im allgemeinen liegen M und $-M^{-1}$ in verschiedenen Ähnlichkeitsklassen der homogenen Modulgruppe, da sie verschiedene Spuren $a+d$ und $-a-d$ haben. In derselben Klasse können sie nur für $a+d=0$ liegen, und tun es auch wirklich, weil dann $M=-M^{-1}$.

Der Ähnlichkeitsklasseneinteilung der Modulgruppe entspricht nun die Einteilung der quadratischen Formen in eigentliche Äquivalenzklassen. Wenn die Koeffizienten der quadratischen Form (37) die Matrix

$$Q = \begin{pmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{pmatrix}$$

bilden, so handelt es sich also um den folgenden

Satz 11. *Wenn der quadratischen Form (oder ihrer Koeffizientenmatrix) Q die Modulsubstitution M (zusammen mit $-M^{-1}$) zugeordnet ist, dann gehört zu $L'QL$ die Matrix $L^{-1}ML$ (zusammen mit $-L^{-1}M^{-1}L$), wo L eine beliebige Modulsubstitution ist und L' die Transponierte von L bedeutet.*

Beweis. Der Satz braucht nur für $L=S$ und $L=T$ bewiesen zu werden.

Es sei

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

dann ist [s. (13a)]

$$S^{-1}MS = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a-c & a-c+b-d \\ c & c+d \end{pmatrix}.$$

Zu M gehört nach (36) die Koeffizientenmatrix

$$Q = \begin{pmatrix} c & \frac{d-a}{2} \\ \frac{d-a}{2} & -b \end{pmatrix},$$

und in der Tat gilt dann

$$S' Q S = \begin{pmatrix} c & c + \frac{d-a}{2} \\ c + \frac{d-a}{2} & c - b + d - a \end{pmatrix} = \begin{pmatrix} c' & \frac{d'-a'}{2} \\ \frac{d'-a'}{2} & -b' \end{pmatrix},$$

wie es sein muß. Ebenso ist [s. (14a)]

$$T^{-1} M T = \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix} = \begin{pmatrix} d & -e \\ -b & a \end{pmatrix}$$

und

$$T' Q T = \begin{pmatrix} -b & \frac{a-d}{2} \\ \frac{a-d}{2} & c \end{pmatrix} = \begin{pmatrix} c'' & \frac{d''-a''}{2} \\ \frac{d''-a''}{2} & -b'' \end{pmatrix},$$

wie es der zu beweisende Satz verlangt.

(Bemerkung. Latimer und MacDuffee [5] und O. Taussky [6] haben eine Zuordnung von Matrizen n-ter Ordnung zu Idealen in algebraischen Körpern n-ten Grades studiert. Wir ziehen hier statt der Ideale in quadratischen Körpern die quadratischen Formen vor, die auch für $\Delta = 0$ brauchbar sind.)

13. Satz 7 zusammen mit Satz 11 ergeben nun den

Satz 12. *Für die quadratischen Formen* $Q(x,y) = A x^2 + B x y + C y^2$ *mit der Diskriminante* $\Delta = B^2 - 4AC = m^2 - 4$ *bildet die Funktion*

$$\Psi\begin{pmatrix} \frac{1}{2}(m-B) & -c \\ A & \frac{1}{2}(m+B) \end{pmatrix}$$

eine Klasseninvariante.

Das folgende Beispiel ist lehrreich: Für $m = 45$, $\Delta = 45^2 - 4 = 43 \cdot 47 = 2021$ gehören die beiden Formen

$$x^2 + xy - 505 y^2 \quad \text{und} \quad 17 x^2 - 7 x y - 29 y^2$$

verschiedenen Klassen an, denn für die erstere findet man

$$\Psi\begin{pmatrix} 22 & 505 \\ 1 & 23 \end{pmatrix} = 42$$

und für die letztere

$$\Psi\begin{pmatrix} 19 & 29 \\ 17 & 26 \end{pmatrix} = -6.$$

Die beiden Formen gehören aber demselben Geschlecht an, weil

$$\left(\frac{A}{43}\right) = +1, \quad \left(\frac{A}{47}\right) = +1$$

für $A = 1$ und $A = 17$ ist[4]). Die Funktion Ψ ist also keine generische Invariante.

[4]) Ein anderes Beispiel dieser Art, für $\Delta = 25^2 - 4 = 621$ wurde mir von Herrn R. Ayoub mitgeteilt.

14. Man kann nun umgekehrt den Satz 12 benutzen, um in vielen Fällen $s(h, k)$ ohne eine auf Gl. (3) gegründete Rekursion zu finden, nämlich zunächst, wenn die Klassenzahl zu $\Delta = m^2 - 4$ gleich 1 ist. Dies trifft nur für $|m| = 0, 1, 3$, also für $\Delta = -4, -3, 5$ zu. Da für diese m alle unimodularen Substitutionen der Spur $a + d = m$ zur gleichen Klasse gehören müssen, so haben wir hier

$$\Psi\begin{pmatrix} a & b \\ c & m-a \end{pmatrix} = \Psi\begin{pmatrix} m & -1 \\ 1 & 0 \end{pmatrix}$$

mit $a(m-a) - bc = 1$, $m = 0, \pm 1, \pm 3$, wobei wir $c > 0$ nehmen können. Wenn man in diese Gleichung die Werte aus (11) einsetzt, so erhält man den

Satz 13. *Es ist für* $c > 0$ *und* $a(m-a) \equiv 1 \pmod{c}$

$$s(a, c) = \frac{(1-c)\,m}{12\,c}$$

für $m = 0, \pm 1, \pm 3$.

Dieser Satz besagt für $m = 0$, also $a^2 \equiv -1 \pmod{c}$ nur einen Teil von Satz 1. Für ungerades m aber ist $a(m-a)$ gerade, also c ungerade. Dann lassen sich die Aussagen für $m = \pm 1, \pm 3$ auch folgendermaßen schreiben:

Für ungerades $c > 0$ und $(2a \pm 1)^2 \equiv -3 \pmod{c}$ ist

$$s(a, c) = \pm \frac{c-1}{12c}; \tag{39}$$

und für $(2a \pm 3)^2 \equiv 5 \pmod{c}$ ist

$$s(a, c) = \pm \frac{c-1}{4c}. \tag{40}$$

Dieser Fall ist besonders bemerkenswert, da er zu $\Delta > 0$ gehört, wo $M(\tau)$ hyperbolisch ist und nur Fixpunkte auf der reellen Achse hat. Ein Beispiel für (40) ist

$$a = 16,\ c = 61,\ m = +3,\ (32+3)^2 = 1225 \equiv 5 \pmod{61},$$

also

$$s(16, 61) = \frac{60}{4 \cdot 61} = \frac{15}{61}.$$

Auch wenn die Klassenzahl von Δ nicht 1 aber klein ist, könnte man Ergebnisse erzielen. Zum Beispiel für $m = 4$, $\Delta = 12$ gibt es zwei Klassen quadratischer Formen, von denen

$$x^2 - 4xy + y^2 \quad \text{und} \quad -x^2 + 4xy - y^2$$

Repräsentanten sind. Diesen Formen sind beziehungsweise die unimodularen Matrizen

$$\begin{pmatrix} 4 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 & 1 \\ -1 & 4 \end{pmatrix}$$

zugeordnet, für die man erhält:

$$\Psi\begin{pmatrix} 4 & -1 \\ 1 & 0 \end{pmatrix} = 1, \quad \Psi\begin{pmatrix} 0 & 1 \\ -1 & 4 \end{pmatrix} = -1.$$

Wir haben also mit $c>0$, $a(4-a)-bc=1$

$$\Psi\begin{pmatrix} a & b \\ c & 4-a \end{pmatrix} = \pm 1$$

oder

$$s(a,c) = \frac{4-3c\mp c}{12c}. \tag{41}$$

Sei zunächst $3\nmid c$. Dann wissen wir nach Hilfssatz 2, daß 3 kein Teiler des Nenners von $s(a,c)$ sein kann, so daß also folgt

$$s(a,c) = \begin{cases} \frac{4-4c}{12c} = \frac{1-c}{3c} & c\equiv 1 \pmod 3, \\ \frac{4-2c}{12c} = \frac{2-c}{6c} & c\equiv 2 \pmod 3. \end{cases} \tag{42}$$

Ist ferner $3/c$, so zeigt die Bedingung $a(4-a)\equiv 1 \pmod c$ oder

$$(a-2)^2 \equiv 3 \pmod c, \tag{43}$$

daß $9\nmid c$, also nur $c\equiv \pm 3 \pmod 9$ in Frage kommen.

Nun folgt aus (41)

$$12c\,s(a,c) \equiv 4\mp c \pmod{3c},$$

anderseits aus (7), Hilfssatz 3,

$$12ac\,s(a,c) \equiv a^2+1 \pmod{3c},$$

also

$$4a\mp ac \equiv a^2+1 \pmod 9.$$

Wegen (43) ist

$$a-2\equiv 0 \pmod 3, \quad (a-2)^2\equiv 0 \pmod 9,$$

also

$$3\mp ac \equiv 0 \pmod 9,$$

oder

$$3\pm c\equiv 0 \pmod 9,$$

so daß, da die oberen und unteren Vorzeichen hier und in (41) sich entsprechen,

$$s(a,c) = \begin{cases} \frac{4-4c}{12c} & c\equiv -3 \pmod 9 \\ \frac{4-2c}{12c} & c\equiv 3 \pmod 9. \end{cases}$$

Zusammenfassend haben wir also

Satz 14. *Wenn* $c>0$, $(a-2)^2\equiv 3 \pmod c$, *so ist*

$$s(a,c) = \begin{cases} \frac{1-c}{3c} & \textit{für } c\equiv 1 \pmod 3 \textit{ und für } c\equiv -3 \pmod 9, \\ \frac{2-c}{6c} & \textit{für } c\equiv -1 \pmod 3 \textit{ und für } c\equiv 3 \pmod 9. \end{cases}$$

Von dieser Art könnte man noch weitere Sätze für geeignetes $\Delta>0$ aufstellen.

15. Es bleibt noch der parabolische Fall $\Delta = 0$, $m = \pm 2$ zu diskutieren. Hier haben wir, wie schon in § 11 erwähnt ist, unendlich viele Klassen.

Sei zunächst

$$a + d = 2, \tag{44}$$

also

$$(d-1)^2 = -bc.$$

Zu

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

gehört die quadratische Form

$$\begin{aligned} Q(x,y) &= cx^2 + 2(d+1)xy - by^2 \\ &= \frac{1}{c}(cx + (d-1)y)^2. \end{aligned}$$

Wenn

$$\lambda = (c, d-1), \tag{45}$$

$$c = \lambda\gamma, \quad d - 1 = \lambda\delta$$

gesetzt wird, erhalten wir

$$Q(x,y) = \frac{\lambda^2}{c}(\gamma x + \delta y)^2.$$

Hier ist einerseits λ^2/c ganz, denn

$$\lambda^2 = (c^2, (d-1)^2)$$

zugleich mit

$$(d-1)^2 \equiv 0 \pmod{c}. \tag{46}$$

Anderseits ist

$$(\gamma, \delta) = 1,$$

also lassen sich α, β finden, so daß

$$\begin{aligned} x' &= \alpha x + \beta y \\ y' &= \gamma x + \delta y \end{aligned}$$

eine Modulsubstitution darstellt, und somit ist

$$Q(x,y) \sim \frac{\lambda^2}{c} y^2.$$

Daher gilt innerhalb der Modulgruppe die Äquivalenz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \sim \begin{pmatrix} 1 & -\frac{\lambda^2}{c} \\ 0 & 1 \end{pmatrix},$$

so daß also

$$\Psi\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \Psi\begin{pmatrix} 1 & -\frac{\lambda^2}{c} \\ 0 & 1 \end{pmatrix} = -\frac{\lambda^2}{c}$$

ist. Ausführlich heißt dies, unter der Annahme $c>0$ (sonst könnte man die Vorzeichen umkehren und $a+d=-2$ betrachten),

$$\frac{2}{c} - 12\,s(a,c) - 3 = -\frac{\lambda^2}{c}, \tag{47}$$

unter der Bedingung (46).

Wenn man endlich statt (44) die Bedingung

$$a+d=-2$$

nimmt, so werden (45), (46) durch

$$\lambda = (c, d+1),$$
$$(d+1)^2 \equiv 0 \pmod{c}$$

ersetzt, und statt (47) erscheint, $c>0$ vorausgesetzt,

$$-\frac{2}{c} - 12\,s(a,c) + 3 = \frac{\lambda^2}{c}.$$

Schließlich beachten wir noch $s(a,c)=s(d,c)$ und erhalten somit den

Satz 15. *Ist* $c>0$, $(d\pm 1)^2\equiv 0 \pmod{c}$, *so ist*

$$s(d,c) = \mp\frac{\lambda^2+2-3c}{12c},$$

mit $\lambda=(c, d\pm 1)$, *wobei die unteren und die oberen Vorzeichen zusammengehören.*

Für quadratfreies c ist stets $\lambda=c$, und der Satz besagt dann nicht mehr als (15). Er enthält jedoch eine neue Aussage für c mit quadratischem Faktor. Wir erhalten z.B. für

$$c=50, \quad d=10j\pm 1, \quad j=1,2,3,4, \quad \text{also} \quad \lambda=10$$
$$s(10j\pm 1, 50) = \pm\frac{100+2-150}{600} = \mp\frac{2}{25}.$$

Literatur

[1] Mordell, L. J.: Lattice points in a tetrahedron and generalized Dedekind sums. J. Indian Math. Soc. **15**, 41—46 (1951). — [2] Rademacher, H.: Generalization of the reciprocity formula for Dedekind sums. Duke Math. J. **21**, 391—397 (1954). — [3] Rademacher, H., and A. Whiteman: Theorems on Dedekind sums. Amer. Math. J. **63**, 377—407 (1941). — [4] Reidemeister, K.: Einführung in die kombinatorische Topologie. Braunschweig 1932. — [5] Latimer, C. G., and C. C. MacDuffee: A correspondence between classes of ideals and classes of matrices. Ann. of Math. (2) **34**, 313—316 (1933). — [6] Taussky, O.: On a theorem of Latimer and MacDuffee. Canad. J. Math. **1**, 300—302 (1949).

Prof. H. Rademacher, University of Pennsylvania, Philadelphia 4, Pa. (USA)

(Eingegangen am 24. August 1955)

Stöhr, A.
Math. Zeitschr. Bd. 63, S. 464—477 (1956)

Zur Funktionentheorie im Raum der symmetrischen Matrizen

Dem Andenken meines verehrten Lehrers I. Schur gewidmet

Von

Alfred Stöhr

Inhalt

Seite
§ 1. Fragestellungen 464
§ 2. Rangerniedrigende und rangerhaltende lineare Funktionen 467
§ 3. Differentiell rangerniedrigende und differentiell spezielle Funktionen 471
§ 4. Die Eigenwerte bei rangerniedrigender linearer Abbildung 473
§ 5. Ein Analogon zum Schwarzschen Lemma 475

§ 1. Fragestellungen

In der Funktionentheorie einer Veränderlichen ist die Gesamtheit der Funktionen, die den Einheitskreis $z\bar{z}<1$ konform auf sich abbilden, bekanntlich durch

$$\eta\frac{z-\alpha}{-\bar{\alpha}z+1} \qquad (|\alpha|<1,\ |\eta|=1) \tag{1}$$

gegeben. Bei jeder solchen Abbildung bleiben nichteuklidische Abstände invariant. Für jede im Einheitskreis reguläre Funktion $f(z)$, die dort der Bedingung $|f(z)|<1$ genügt, die also den Einheitskreis in sich abbildet, gilt dann für je zwei Punkte z, z_1 aus dem Einheitskreis die Ungleichung

$$\delta\big(f(z), f(z_1)\big)\leqq\delta(z, z_1), \tag{2}$$

wo δ den nichteuklidischen Abstand bezeichnet. Für den Nachweis dieser Tatsache kann man sich, indem man Hilfsabbildungen von der Gestalt (1) heranzieht, ohne Beschränkung der Allgemeinheit auf den Spezialfall $z_1=f(z_1)=0$ beschränken. Da die nichteuklidische Entfernung $\delta(z, 0)$ eine monotone Funktion von $|z|$ ist, läuft in diesem Spezialfall der Nachweis von (2) auf das Schwarzsche Lemma

$$|f(z)|\leqq|z| \tag{3}$$

hinaus. Aus (3) folgt

$$|f'(0)|\leqq 1, \tag{4}$$

woraus umgekehrt nach Anwendung von Transformationen (1) die Ungleichung

$$\frac{|df(z)|}{1-|f(z)|^2}\leqq\frac{|dz|}{1-|z|^2}$$

und daraus durch Integration wieder (2) und (3) hergeleitet werden können.

Ziel der folgenden Untersuchungen ist es, die genannten Sätze auf gewisse Funktionen auszudehnen, deren unabhängige und abhängige Veränderliche n-reihige bzw. m-reihige symmetrische Matrizen sind; die Elemente der abhängigen Matrix sollen dabei analytisch von den Elementen der unabhängigen Matrix abhängen. Beispiele solcher Funktionen erhält man für $m = n$ wie folgt: $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ seien n-reihige quadratische Matrizen. $\begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}$ heiße symplektisch, wenn $\mathfrak{A}'\mathfrak{C} = \mathfrak{C}'\mathfrak{A}$, $\mathfrak{B}'\mathfrak{D} = \mathfrak{D}'\mathfrak{B}$, $\mathfrak{A}'\mathfrak{D} - \mathfrak{C}'\mathfrak{B} = \mathfrak{E}$ ist. Dann führt die Abbildung

$$\mathfrak{Z} \to \mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} \quad (\text{für } \det(\mathfrak{C}\mathfrak{Z} + \mathfrak{D}) \neq 0) \tag{5}$$

symmetrische Matrizen $\mathfrak{Z}$ in symmetrische Matrizen $\mathfrak{W}$ über; wir nennen $\mathfrak{W}$ eine symplektische linear-gebrochene Funktion von $\mathfrak{Z}$.

Sind $\mathfrak{Z}_1, \mathfrak{Z}_2, \mathfrak{Z}_3, \mathfrak{Z}_4$ irgend vier symmetrische Matrizen, so bilde man mit ihnen den Ausdruck

$$\mathfrak{R} = \mathfrak{R}(\mathfrak{Z}_1, \mathfrak{Z}_2, \mathfrak{Z}_3, \mathfrak{Z}_4) = (\mathfrak{Z}_1 - \mathfrak{Z}_3)(\mathfrak{Z}_1 - \mathfrak{Z}_4)^{-1}(\mathfrak{Z}_2 - \mathfrak{Z}_4)(\mathfrak{Z}_2 - \mathfrak{Z}_3)^{-1}$$

(falls die rechte Seite einen Sinn hat). Es empfiehlt sich, nicht nur diese Matrix, sondern auch jede der gleichen Ähnlichkeitsklasse angehörige Matrix $\mathfrak{M}\mathfrak{R}\mathfrak{M}^{-1}$ als Doppelverhältnis von $\mathfrak{Z}_1, \mathfrak{Z}_2, \mathfrak{Z}_3, \mathfrak{Z}_4$ zu bezeichnen; die (bekanntlich nur von der Ähnlichkeitsklasse abhängenden) Eigenwerte von $\mathfrak{R}$ nennen wir die Eigenwerte des Doppelverhältnisses[1]). Die durch $\mathfrak{R}$ repräsentierte Ähnlichkeitsklasse ist gegenüber symplektischen linear-gebrochenen Substitutionen invariant, was man durch mehrmalige Anwendung der Formel

$$\begin{aligned} \mathfrak{W}_i - \mathfrak{W}_k &= (\mathfrak{A}\mathfrak{Z}_i + \mathfrak{B})(\mathfrak{C}\mathfrak{Z}_i + \mathfrak{D})^{-1} - (\mathfrak{A}\mathfrak{Z}_k + \mathfrak{B})(\mathfrak{C}\mathfrak{Z}_k + \mathfrak{D})^{-1} \\ &= (\mathfrak{A}\mathfrak{Z}_i + \mathfrak{B})(\mathfrak{C}\mathfrak{Z}_i + \mathfrak{D})^{-1} - (\mathfrak{Z}_k\mathfrak{C}' + \mathfrak{D}')^{-1}(\mathfrak{Z}_k\mathfrak{A}' + \mathfrak{B}') \\ &= (\mathfrak{C}\mathfrak{Z}_k + \mathfrak{D})'^{-1}(\mathfrak{Z}_i - \mathfrak{Z}_k)(\mathfrak{C}\mathfrak{Z}_i + \mathfrak{D})^{-1} \end{aligned}$$

erkennt.

Die folgenden Definitionen und Sätze entnehmen wir einer Arbeit von Siegel [2]. — Im Raum der symmetrischen Matrizen $\mathfrak{Z}$ werde der Einheitskreis durch

$$\mathfrak{E} - \mathfrak{Z}\overline{\mathfrak{Z}} > 0 \tag{6}$$

definiert. Für zwei im Einheitskreis gelegene Matrizen $\mathfrak{Z}, \mathfrak{Z}_1$ bilde man

$$\mathfrak{R} = \mathfrak{R}(\mathfrak{Z}, \overline{\mathfrak{Z}}^{-1}, \mathfrak{Z}_1, \overline{\mathfrak{Z}}_1^{-1}) = (\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{Z} - \overline{\mathfrak{Z}}_1^{-1})^{-1}(\overline{\mathfrak{Z}}^{-1} - \overline{\mathfrak{Z}}_1^{-1})(\overline{\mathfrak{Z}}^{-1} - \mathfrak{Z}_1)^{-1}.$$

Durch die Umformung

$$\mathfrak{R} = (\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{E} - \overline{\mathfrak{Z}}_1\mathfrak{Z})^{-1}(\overline{\mathfrak{Z}} - \overline{\mathfrak{Z}}_1)(\mathfrak{E} - \mathfrak{Z}_1\overline{\mathfrak{Z}})^{-1} \tag{7}$$

wird dieser Ausdruck auch für Matrizen mit verschwindender Determinante sinnvoll. [Für $\mathfrak{Z}_1 = 0$ nimmt (7) die einfache Gestalt $\mathfrak{R} = \mathfrak{Z}\overline{\mathfrak{Z}}$ an.] Zu (7)

[1]) Über Doppelverhältnisse vgl. auch [1] sowie die dort angegebene Literatur.

bestimme man eine Matrix $\mathfrak{T}$ mittels der konvergenten Reihe

$$\mathfrak{T} = 4\mathfrak{R}\left(\sum_{k=0}^{\infty} \frac{\mathfrak{R}^k}{2k+1}\right)^2$$

(also $\mathfrak{T}\mathfrak{g}\frac{\sqrt{\mathfrak{T}}}{2} = \sqrt{\mathfrak{R}}$) und setze dann $s = \sqrt{\text{Spur}\,\mathfrak{T}}$. Die dadurch definierte nichtnegative reelle Zahl s wollen wir hier als die symplektische Entfernung der Matrizen $\mathfrak{Z}$, $\mathfrak{Z}_1$ im Einheitskreis bezeichnen. Ist ein im Einheitskreis (6) gelegenes rektifizierbares Kurvenstück gegeben, so nennen wir das darüber erstreckte Integral

$$\int ds = 2\int \left| \text{Spur}\,\{d\mathfrak{Z}(\mathfrak{E} - \overline{\mathfrak{Z}}\,\mathfrak{Z})^{-1}\, d\overline{\mathfrak{Z}}\,(\mathfrak{E} - \mathfrak{Z}\,\overline{\mathfrak{Z}})^{-1}\}\right| \tag{8}$$

die symplektische Bogenlänge im Einheitskreis. Jede analytische Abbildung des Einheitskreises (6) auf sich wird durch eine symplektische linear-gebrochene Funktion bewirkt. Bei jeder solchen Abbildung bleibt die durch (7) repräsentierte Ähnlichkeitsklasse und folglich auch die symplektische Entfernung s und die symplektische Bogenlänge (8) invariant. Zu je zwei Punkten $\mathfrak{Z}_1$, $\mathfrak{W}_1$ im Einheitskreis gibt es eine derartige Abbildung, die $\mathfrak{Z}_1$ in $\mathfrak{W}_1$ überführt. Die Gesamtheit aller Abbildungen, die den Einheitskreis festlassen und die Nullmatrix in die Nullmatrix überführen, wird durch

$$\mathfrak{W} = \mathfrak{Z}[\mathfrak{U}] \qquad (\mathfrak{U} \text{ unitär})$$

gegeben[2]). Jede dem Einheitskreis angehörende symmetrische Matrix $\mathfrak{M}$ kann in der Gestalt

$$\mathfrak{M} = \mathfrak{P}[\mathfrak{U}] \qquad (\mathfrak{U} \text{ unitär, } \mathfrak{P} \text{ Diagonalmatrix}) \tag{9}$$

geschrieben werden, wobei die Diagonalelemente $p_1, \ldots, p_n$ von $\mathfrak{P}$ beliebig gewählt werden können bis auf die Einschränkung, daß $|p_1|^2, \ldots, |p_n|^2$ die Eigenwerte der hermiteschen Matrix $\mathfrak{M}\overline{\mathfrak{M}}$ sein müssen. — Soweit die Anknüpfung an die erwähnte Arbeit von SIEGEL.

Unter den analytischen Funktionen $\mathfrak{W} = f(\mathfrak{Z})$ werden wir in § 3 diejenigen als „differentiell rangerniedrigend" aussondern, bei denen stets Rang $d\mathfrak{W} \leq$ Rang $d\mathfrak{Z}$ ist; zu dieser Klasse von Funktionen gehören insbesondere die symplektischen linear-gebrochenen Funktionen. Eines unserer Ergebnisse wird dann lauten, daß bei Abbildungen $\mathfrak{Z} \to \mathfrak{W} = f(\mathfrak{Z})$, die durch eine differentiell rangerniedrigende Funktion bewirkt werden und die den Einheitskreis des $\mathfrak{Z}$-Raumes in den Einheitskreis des $\mathfrak{W}$-Raumes hinein abbilden, alle symplektischen Entfernungen nicht zunehmen. Der Beweis erfolgt so, daß der Satz zunächst für unendlich kleine Entfernungen bewiesen wird; für diesen Fall wird in § 4 noch eine schärfere, die Eigenwerte des Doppelverhältnisses betreffende Aussage hergeleitet.

Solche Funktionen im Raum der symmetrischen Matrizen sind für gewisse elektrotechnische Probleme von Interesse. In der Theorie der linearen Schaltungen wird der Widerstand eines passiven $2n$-Pols durch eine n-reihige

[2]) Wir benutzen die Abkürzung $\mathfrak{A}[\mathfrak{B}] = \mathfrak{B}'\mathfrak{A}\,\mathfrak{B}$.

quadratische Matrix dargestellt. Durch eine passive lineare Schaltung mit $2m$ Eingangsklemmen und $2n$ Ausgangsklemmen wird der $2n$-Pol in einen $2m$-Pol transformiert. In einigen wichtigen Fällen bedeutet das für die Widerstände eine Abbildung von symmetrischen n-reihigen Matrizen mit positivem Realteil auf symmetrische m-reihige Matrizen mit positivem Realteil, die überdies durch eine differentiell rangerniedrigende Funktion bewirkt wird. Die symmetrischen Matrizen $\mathfrak{M}$ mit positivem Realteil lassen sich aber durch die symplektische linear-gebrochene Funktion

$$\mathfrak{Z} = (\mathfrak{M} - \mathfrak{E})\,(\mathfrak{M} + \mathfrak{E})^{-1}$$

umkehrbar eindeutig auf die Matrizen $\mathfrak{Z}$ im Einheitskreis abbilden. Bei dieser Abbildung bleiben die [in der oberen Halbebene mittels

$$\mathfrak{R} = (\mathfrak{M} - \mathfrak{M}_1)\,(\mathfrak{M} + \overline{\mathfrak{M}_1})^{-1}\,(\overline{\mathfrak{M}} - \overline{\mathfrak{M}_1})\,(\overline{\mathfrak{M}} + \mathfrak{M}_1)^{-1}$$

zu definierenden] symplektischen Entfernungen erhalten, so daß die dafür im Einheitskreis gewonnenen Sätze auch auf die Matrizen mit positivem Realteil Anwendung finden können.

§ 2. Rangerniedrigende und rangerhaltende lineare Funktionen

Die Buchstaben $\mathfrak{W}$ und $\mathfrak{Z}$ bezeichnen im folgenden stets symmetrische Matrizen; $\mathfrak{W}$ sei dabei m-reihig, $\mathfrak{Z}$ sei n-reihig.

Die Begriffe „rangerniedrigend" und „rangerhaltend" mögen zunächst für lineare homogene Funktionen erklärt werden. Die Elemente der Matrix $\mathfrak{W}$ mögen linear und homogen von den Elementen der Matrix $\mathfrak{Z}$ abhängen. Die Funktion $\mathfrak{W} = f(\mathfrak{Z})$ und die dadurch bewirkte Abbildung $\mathfrak{Z} \to \mathfrak{W}$ mögen *rangerniedrigend* heißen, wenn stets Rang $\mathfrak{W} \leq$ Rang $\mathfrak{Z}$ ist; sie mögen *rangerhaltend* heißen, wenn sogar stets Rang $\mathfrak{W} =$ Rang $\mathfrak{Z}$ ist[3]). Daß nicht jede lineare homogene Funktion rangerniedrigend ist, zeigt das Beispiel

$$\mathfrak{Z} = \begin{pmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{pmatrix}, \qquad f(\mathfrak{Z}) = \begin{pmatrix} z_{11} & 0 \\ 0 & z_{11} \end{pmatrix}.$$

Eine Funktion, die in der Gestalt $f(\mathfrak{Z}) = \mathfrak{Z}[\mathfrak{A}]$ mit von $\mathfrak{Z}$ unabhängiger Matrix $\mathfrak{A}$ geschrieben werden kann, heiße eine *spezielle* lineare homogene Funktion.

Satz 1. *Jede spezielle lineare homogene Funktion* $\mathfrak{Z}[\mathfrak{A}]$ *ist rangerniedrigend; sie ist rangerhaltend genau dann, wenn Rang* $\mathfrak{A} = n$ *ist.*

Beweis. Es ist Rang $\mathfrak{A}'\mathfrak{Z}\mathfrak{A} \leq$ Rang $\mathfrak{Z}$, woraus die Rangerniedrigung folgt. Der Rang von $\mathfrak{A}$ sei k. Dann ist $\mathfrak{A}$ in ein Produkt $\mathfrak{A} = \mathfrak{D}\mathfrak{C}\mathfrak{B}$ zerlegbar, wo $\mathfrak{B}$ eine m-reihige quadratische Matrix mit von Null verschiedener Determinante, $\mathfrak{D}$ eine n-reihige quadratische Matrix mit von Null verschiedener Determinante und $\mathfrak{C} = (c_{ij})$ eine Matrix ist, deren Elemente sämtlich verschwinden

[3]) „rangerniedrigend" und „Rangerniedrigung" schließen also stets die Möglichkeit ein, daß der Rang erhalten bleibt.

mit Ausnahme von $c_{11}=c_{22}=\cdots=c_{kk}=1$. Die Abbildung $\mathfrak{Z}\to\mathfrak{W}=\mathfrak{Z}[\mathfrak{A}]$ läßt sich dann in die Hintereinanderausführung der drei Abbildungen $\mathfrak{Z}\to\mathfrak{Z}_1=\mathfrak{Z}[\mathfrak{D}]$, $\mathfrak{Z}_1\to\mathfrak{Z}_2=\mathfrak{Z}_1[\mathfrak{C}]$, $\mathfrak{Z}_2\to\mathfrak{W}=\mathfrak{Z}_2[\mathfrak{B}]$ zerlegen. Die erste und dritte dieser Abbildungen sind umkehrbar und rangerhaltend und können als bloße Änderungen eines Koordinatensystems angesehen werden, die zweite Abbildung dagegen führt die n-reihige symmetrische Matrix

$$\begin{pmatrix} z_{11} & \cdots & z_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ z_{n1} & \cdots & z_{nn} \end{pmatrix}$$

in die m-reihige symmetrische Matrix

$$\begin{pmatrix} z_{11} & \cdots & z_{1k} & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ z_{k1} & \cdots & z_{kk} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

über; daraus lassen sich alle behaupteten Aussagen ablesen. – Über die in Satz 1 genannten Aussagen hinaus zeigt dieser Beweis, wie der Raum aller symmetrischen Matrizen $\mathfrak{Z}$ auf einen gewissen linearen Teilraum der symmetrischen Matrizen $\mathfrak{W}$ abgebildet wird.

Nicht jede rangerniedrigende lineare homogene Funktion ist speziell; Gegenbeispiel:

$$\mathfrak{Z}=\begin{pmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{pmatrix}, \qquad f(\mathfrak{Z})=\begin{pmatrix} z_{12} & 0 \\ 0 & 0 \end{pmatrix}.$$

Der folgende Satz zeigt jedoch, daß jede rangerhaltende lineare homogene Funktion speziell ist.

Satz 2. *Die lineare homogene Funktion $f(\mathfrak{Z})$ sei rangerniedrigend, und für mindestens eine feste Matrix $\mathfrak{Z}^{(0)}$ sei Rang $\mathfrak{Z}^{(0)}=$ Rang $f(\mathfrak{Z}^{(0)})=n$.* (Für rangerhaltende Funktionen sind diese Voraussetzungen gewiß erfüllt.) *Dann ist $f(\mathfrak{Z})$ rangerhaltend und speziell.*

Beweis. Man bestimme eine quadratische Matrix $\mathfrak{A}_0$ mit $\mathfrak{A}_0'\mathfrak{A}_0=\mathfrak{Z}^{(0)}$; dies ist möglich, und es ist $\det\mathfrak{A}_0\neq 0$. Man setze $f_1(\mathfrak{Z})=f(\mathfrak{Z}[\mathfrak{A}_0])$; dann ist $f_1(\mathfrak{Z})$ linear homogen und rangerniedrigend, und es ist Rang $f_1(\mathfrak{E})=n$. (In diesem Beweis soll $\mathfrak{E}$ immer die n-reihige Einheitsmatrix bedeuten.) $\mathfrak{Z}_k$ sei für $k=1,\ldots,n$ die Matrix, für die $z_{kk}=1$ ist und alle übrigen Elemente verschwinden. Da f_1 rangerniedrigend ist, hat $f_1(\mathfrak{Z}_k)$ höchstens den Rang 1. Wir bestimmen m-spaltige Zeilen $\mathfrak{a}_k$ mit $f_1(\mathfrak{Z}_k)=\mathfrak{a}_k'\mathfrak{a}_k$ und stellen

diese Zeilen zu einer n-zeiligen, m-spaltigen Matrix $\mathfrak{A}_1 = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_n \end{pmatrix}$ zusammen. Dann ist $f_1(\mathfrak{Z}_k) = \mathfrak{Z}_k[\mathfrak{A}_1]$ und folglich wegen der Linearität von f_1 auch $f_1(\mathfrak{Z}) = \mathfrak{Z}[\mathfrak{A}_1]$ für jede Diagonalmatrix $\mathfrak{Z}$. Durch Einsetzen von $\mathfrak{Z} = \mathfrak{E}$ ergibt sich Rang $\mathfrak{A}_1' \mathfrak{A}_1 = n$, woraus Rang $\mathfrak{A}_1 = n$ folgt. Wir ergänzen $\mathfrak{A}_1$ durch Hinzufügen von $m - n$ weiteren Zeilen zu einer quadratischen Matrix $\mathfrak{A}_2 = \begin{pmatrix} \mathfrak{A}_1 \\ * \end{pmatrix}$ dergestalt, daß $\mathfrak{A}_2$ eine Inverse besitzt. Dann ist $(\mathfrak{E}\ 0)\mathfrak{A}_2 = \mathfrak{A}_1$, also $\mathfrak{A}_1 \mathfrak{A}_2^{-1} = (\mathfrak{E}\ 0)$. Wir setzen $f_1(\mathfrak{Z})[\mathfrak{A}_2^{-1}] = f_2(\mathfrak{Z})$; diese Funktion ist wieder linear homogen und rangerniedrigend. Wir zerlegen für beliebiges $\mathfrak{Z}$

$$f_2(\mathfrak{Z}) = \begin{pmatrix} f_3(\mathfrak{Z}) & * \\ * & * \end{pmatrix},$$

wo $f_3(\mathfrak{Z})$ eine n-reihige quadratische Matrix bedeutet; auch die Funktion f_3 ist linear homogen und rangerniedrigend. Speziell für Diagonalmatrizen $\mathfrak{Z}$ ist dann

$$f_2(\mathfrak{Z}) = \mathfrak{Z}[\mathfrak{A}_1 \mathfrak{A}_2^{-1}] = \mathfrak{Z}[(\mathfrak{E}\ 0)] = \begin{pmatrix} \mathfrak{Z} & 0 \\ 0 & 0 \end{pmatrix},$$

also $f_3(\mathfrak{Z}) = \mathfrak{Z}$. Für beliebige Matrizen $\mathfrak{Z}$ gilt wegen der Linearität der Funktion f_3

$$f_3(\lambda \mathfrak{E} - \mathfrak{Z}) = \lambda \mathfrak{E} - f_3(\mathfrak{Z}),$$

und da f_3 rangerniedrigend ist, folgt für festes λ aus $\det(\lambda \mathfrak{E} - \mathfrak{Z}) = 0$ die Gleichung $\det(\lambda \mathfrak{E} - f_3(\mathfrak{Z})) = 0$. Falls die Eigenwerte von $\mathfrak{Z}$ paarweise verschieden sind, haben also $\det(\lambda \mathfrak{E} - \mathfrak{Z})$ und $\det(\lambda \mathfrak{E} - f_3(\mathfrak{Z}))$ als Polynome in λ die gleichen Nullstellen und die gleichen höchsten Koeffizienten. Die daraus folgende Gleichung $\det(\lambda \mathfrak{E} - \mathfrak{Z}) = \det(\lambda \mathfrak{E} - f_3(\mathfrak{Z}))$ ist eine Identität in λ und den Elementen von $\mathfrak{Z}$ und gilt deshalb ohne Einschränkung. Insbesondere ist $\det \mathfrak{Z} = \det f_3(\mathfrak{Z})$. Man setze $\mathfrak{Z}_{ik}$ gleich der Matrix, in der $z_{ik} = z_{ki} = 1$ ist und alle übrigen Elemente verschwinden, und verstehe unter

$$\mathfrak{Z}_0 = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix}$$

eine Diagonalmatrix mit unbestimmten Elementen. Aus

$$\det(\mathfrak{Z}_0 + \mathfrak{Z}_{ik}) = \det f_3(\mathfrak{Z}_0 + \mathfrak{Z}_{ik}) = \det(\mathfrak{Z}_0 + f_3(\mathfrak{Z}_{ik}))$$

folgt durch Vergleich der Koeffizienten der Produkte der $z_1, \ldots, z_n$, daß die symmetrische Matrix $f_3(\mathfrak{Z}_{ik})$ gleich $\pm \mathfrak{Z}_{ik}$ ist. $\mathfrak{A}_3$ sei eine Diagonalmatrix mit Diagonalelementen ± 1. Man setze $f_4(\mathfrak{Z}) = f_3(\mathfrak{Z})[\mathfrak{A}_3]$. Diese Funktion f_4 besitzt alle bei der Funktion f_3 konstatierten Eigenschaften, und durch eine stets mögliche passende Wahl der Vorzeichen in der Diagonale von $\mathfrak{A}_3$ sei für $2 \leqq k \leqq n$ ferner erreicht, daß $f_4(\mathfrak{Z}_{1k}) = +\mathfrak{Z}_{1k}$ ist. Sei nun $2 \leqq i, k \leqq n$ und $i \neq k$. Vergleicht man in

$$\begin{aligned} \det(\mathfrak{Z}_0 + \mathfrak{Z}_{1i} + \mathfrak{Z}_{1k} + \mathfrak{Z}_{ik}) &= \det f_4(\mathfrak{Z}_0 + \mathfrak{Z}_{1i} + \mathfrak{Z}_{1k} + \mathfrak{Z}_{ik}) \\ &= \det(\mathfrak{Z}_0 + \mathfrak{Z}_{1i} + \mathfrak{Z}_{1k} + f_4(\mathfrak{Z}_{ik})) \end{aligned}$$

wieder die Koeffizienten der Produkte der $z_1, \ldots, z_n$, so erkennt man, daß jetzt $f_4(\mathfrak{Z}_{ik}) = +\mathfrak{Z}_{ik}$ ist; wegen der Linearität von f_4 gilt also $f_4(\mathfrak{Z}) = \mathfrak{Z}$ für alle $\mathfrak{Z}$. Wir bezeichnen mit $\mathfrak{E}^*$ die $(m-n)$-reihige Einheitsmatrix, setzen $\mathfrak{A}_4 = \begin{pmatrix} \mathfrak{A}_3 & 0 \\ 0 & \mathfrak{E}^* \end{pmatrix}$ und betrachten die lineare homogene Funktion

$$f_5(\mathfrak{Z}) = f_2(\mathfrak{Z})\,[\mathfrak{A}_4] = \begin{pmatrix} \mathfrak{Z} & f_6(\mathfrak{Z}) \\ (f_6(\mathfrak{Z}))' & f_7(\mathfrak{Z}) \end{pmatrix}. \tag{10}$$

Weil $\det \mathfrak{A}_4 \neq 0$ und $f_2(\mathfrak{Z})$ rangerniedrigend ist, ist Rang $f_5(\mathfrak{Z}) \leqq$ Rang $\mathfrak{Z}$; andererseits folgt aus der Gestalt der rechten Seite von (10) Rang $f_5(\mathfrak{Z}) \geqq$ Rang $\mathfrak{Z}$. Die Funktion $f_5(\mathfrak{Z})$ ist demnach rangerhaltend; also sind auch $f_2(\mathfrak{Z})$, $f_1(\mathfrak{Z})$ und schließlich auch das ursprüngliche $f(\mathfrak{Z})$ rangerhaltend. Für Diagonalmatrizen $\mathfrak{Z}$ verschwinden $f_6(\mathfrak{Z})$ und $f_7(\mathfrak{Z})$. Wir zeigen, daß $f_6(\mathfrak{Z})$ identisch verschwindet; dazu genügt wegen der Linearität von f_6 der Nachweis, daß $f_6(\mathfrak{Z}_{ik})$ für $1 \leqq i, k \leqq n$ und $i \neq k$ verschwindet. Die Matrizen $\mathfrak{Z}^{(1)} = \mathfrak{Z}_i + \mathfrak{Z}_k + \mathfrak{Z}_{ik}$ und $\mathfrak{Z}^{(2)} = -\mathfrak{Z}_i - \mathfrak{Z}_k + \mathfrak{Z}_{ik}$ haben den Rang 1; es ist also für $j = 1, 2$ auch

$$\text{Rang } f_5(\mathfrak{Z}^{(j)}) = \text{Rang} \begin{pmatrix} \mathfrak{Z}^{(j)} & f_6(\mathfrak{Z}_{ik}) \\ * & * \end{pmatrix} = 1.$$

Die Spalten von $f_6(\mathfrak{Z}_{ik})$ müssen also sowohl zu den Spalten von $\mathfrak{Z}^{(1)}$ als auch zu den Spalten von $\mathfrak{Z}^{(2)}$ proportional sein, woraus ihr Verschwinden folgt. Also ist für alle $\mathfrak{Z}$

$$f_5(\mathfrak{Z}) = \begin{pmatrix} \mathfrak{Z} & 0 \\ 0 & f_7(\mathfrak{Z}) \end{pmatrix};$$

da f_5 rangerniedrigend ist, folgt daraus $f_7(\mathfrak{Z}) = 0$. Wir haben nun

$$f(\mathfrak{Z}\,[\mathfrak{A}_0])\,[\mathfrak{A}_2^{-1}\mathfrak{A}_4] = f_5(\mathfrak{Z}) = \mathfrak{Z}\,[(\mathfrak{E}\,0)],$$

woraus $f(\mathfrak{Z}) = \mathfrak{Z}\,[\mathfrak{A}]$ mit $\mathfrak{A} = \mathfrak{A}_0^{-1}(\mathfrak{E}\,0)\mathfrak{A}_4^{-1}\mathfrak{A}_2$ folgt. $f(\mathfrak{Z})$ ist also speziell.

Satz 3. *Ist $f(\mathfrak{Z}) = \mathfrak{Z}\,[\mathfrak{A}]$ eine spezielle lineare homogene Funktion, so ist durch $f(\mathfrak{Z})$ die Matrix $\mathfrak{A}$ bis auf einen willkürlich bleibenden skalaren Faktor ± 1 eindeutig bestimmt.*

Beweis. Aus $\mathfrak{A}'\mathfrak{Z}\mathfrak{A} = \mathfrak{B}'\mathfrak{Z}\mathfrak{B}$ für alle $\mathfrak{Z}$ haben wir $\mathfrak{A} = \pm\mathfrak{B}$ zu schließen. $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ seien die Zeilen von $\mathfrak{A}$, und $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ seien die Zeilen von $\mathfrak{B}$, ferner mögen $\mathfrak{Z}_k$ und $\mathfrak{Z}_{ik}$ dieselbe Bedeutung wie im Beweis von Satz 2 haben. $\mathfrak{A}'\mathfrak{Z}_k\mathfrak{A} = \mathfrak{B}'\mathfrak{Z}_k\mathfrak{B}$ ergibt $\mathfrak{a}_k'\mathfrak{a}_k = \mathfrak{b}_k'\mathfrak{b}_k$, also $\mathfrak{a}_k = \pm\mathfrak{b}_k$. Angenommen, in dieser Gleichung gälte nicht für alle Indizes k das gleiche Vorzeichen; etwa für die Indizes i, k sei $\mathfrak{a}_i = +\mathfrak{b}_i$ und $\mathfrak{a}_k = -\mathfrak{b}_k$; überdies sei $\mathfrak{a}_i \neq 0$ und $\mathfrak{a}_k \neq 0$ (da sonst das betreffende Vorzeichen sowieso irrelevant wäre). Aus $\mathfrak{A}'\mathfrak{Z}_{ik}\mathfrak{A} = \mathfrak{B}'\mathfrak{Z}_{ik}\mathfrak{B}$ folgte dann $\mathfrak{a}_i'\mathfrak{a}_k + \mathfrak{a}_k'\mathfrak{a}_i = \mathfrak{b}_i'\mathfrak{b}_k + \mathfrak{b}_k'\mathfrak{b}_i = \mathfrak{a}_i'(-\mathfrak{a}_k) + (-\mathfrak{a}_k)'\mathfrak{a}_i$, also $\mathfrak{a}_i'\mathfrak{a}_k + \mathfrak{a}_k'\mathfrak{a}_i = 0$. Eine solche Gleichung kann nur bestehen, wenn die Zeilen $\mathfrak{a}_i$ und $\mathfrak{a}_k$ proportional sind, also etwa $\mathfrak{a}_i = \lambda\mathfrak{a}_k$. Das gibt $2\lambda\,\mathfrak{a}_k'\mathfrak{a}_k = \mathfrak{a}_i'\mathfrak{a}_k + \mathfrak{a}_k'\mathfrak{a}_i = 0$. Da die Matrix $\mathfrak{a}_k'\mathfrak{a}_k$ nicht die Nullmatrix ist, folgt $\lambda = 0$, $\mathfrak{a}_i = 0$ entgegen der Annahme.

§ 3. Differentiell rangerniedrigende und differentiell spezielle Funktionen.

Wir betrachten nun beliebige differenzierbare Funktionen $\mathfrak{W} = f(\mathfrak{Z})$; darunter soll verstanden werden, daß die Elemente von $\mathfrak{W}$ total differenzierbare Funktionen der Elemente von $\mathfrak{Z}$ sind. Die Matrix der Differentiale $d\mathfrak{W}$ ist dann eine lineare homogene Funktion der Matrix der Differentiale $d\mathfrak{Z}$. Wir nennen $\mathfrak{W} = f(\mathfrak{Z})$ eine *differentiell rangerniedrigende* bzw. *differentiell spezielle* bzw. *differentiell rangerhaltende* Funktion, wenn $d\mathfrak{W}$ für alle in Betracht kommenden $\mathfrak{Z}$ eine rangerniedrigende bzw. spezielle bzw. rangerhaltende Funktion von $d\mathfrak{Z}$ ist. Die Hintereinanderausführung differentiell rangerniedrigender bzw. differentiell spezieller bzw. differentiell rangerhaltender Funktionen ergibt, wenn sie ausführbar ist, wieder eine solche Funktion; im Falle $m = n$ ist für differentiell rangerhaltende Funktionen eine Umkehrfunktion vorhanden, und diese ist wieder differentiell rangerhaltend. — Für beliebige differentiell spezielle Funktionen könnte man die in der Relation

$$d\mathfrak{W} = (d\mathfrak{Z})\,[\mathfrak{M}] \tag{11}$$

auftretende, bis auf ihr Vorzeichen bestimmte und im allgemeinen von $\mathfrak{Z}$ abhängende Matrix $\mathfrak{M}$ die symmetrische Ableitung von $f(\mathfrak{Z})$ nennen.

Wir wollen uns noch durch einige Beispiele überzeugen, daß es differentiell spezielle Funktionen gibt[4]). Ein triviales Beispiel solcher Funktionen sind die speziellen linearen homogenen Funktionen; ein weiteres wichtiges Beispiel wird durch die symplektische gebrochen lineare Funktion

$$\mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})\,(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$$

geliefert, für diese gilt nämlich $d\mathfrak{W} = (d\mathfrak{Z})\,[(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}]$. Die Hintereinanderausführung der in diesen beiden Beispielen auftretenden Funktionen liefert eine Klasse von Funktionen, die in der Elektrotechnik bei der Transformation von symmetrischen Widerstandsmatrizen auftreten; man vergleiche die Bemerkungen am Ende von § 1.

Der Rest dieses Paragraphen bezieht sich auf den Spezialfall $m = 1$. Die Matrix $\mathfrak{W}$ besteht dann nur aus einem Element, das eine Funktion $w = w(\mathfrak{Z})$ ist; die Matrix $\mathfrak{M} = (\mu_1, \ldots, \mu_n)$ besteht nur aus einer Zeile, und die $\mu_1, \ldots, \mu_n$ sind im allgemeinen ebenfalls Funktionen von $\mathfrak{Z}$. Nach (11) sind die Gleichungen

$$\frac{\partial w}{\partial z_{ik}} = \varepsilon_{ik}\,\mu_i\,\mu_k \qquad (i = 1, \ldots, n;\ k = 1, \ldots, n) \tag{12}$$

zu erfüllen, abkürzend ist

$$\varepsilon_{ik} = \begin{cases} 1 & \text{für } i = k, \\ 2 & \text{für } i \neq k \end{cases}$$

[4]) Die Frage nach der Aufstellung aller dieser Funktionen, d.h. nach der Bestimmung aller Lösungen des Systems partieller Differentialgleichungen (11), bleibt offen.

gesetzt. Aus (12) läßt sich leicht ein System von Differentialgleichungen herleiten, in denen $\mu_1, \ldots, \mu_n$ nicht mehr vorkommen:

$$\frac{1}{\varepsilon_{ik}\varepsilon_{jl}}\frac{\partial w}{\partial z_{ik}}\frac{\partial w}{\partial z_{jl}} = \frac{1}{\varepsilon_{ij}\varepsilon_{kl}}\frac{\partial w}{\partial z_{ij}}\frac{\partial w}{\partial z_{kl}}, \tag{13}$$

sowie ein System von Integrabilitätsbedingungen:

$$\varepsilon_{ik}\frac{\partial(\mu_i\mu_k)}{\partial z_{jl}} = \varepsilon_{jl}\frac{\partial(\mu_j\mu_l)}{\partial z_{ik}}\ ^{5}).$$

Da einige Lösungen der Gln. (12) bzw. (13) bekannt sind, die einige Parameter enthalten (z.B. $w = \mathfrak{Z}[\mathfrak{M}]$ mit konstantem $\mathfrak{M} = (\mu_1, \ldots, \mu_n)$, wo $\mu_1, \ldots, \mu_n$ als Parameter aufgefaßt werden mögen), so kann man aus diesen in üblicher Weise durch Differentiations- und Eliminationsprozesse weitere Lösungen des Gleichungssystems (13) gewinnen. Man kann zur Auffindung weiterer Lösungen von (12), (13) auch folgendermaßen vorgehen: $\varphi_{ik}(w)$ $(i = 1, \ldots, n;\ k = 1, \ldots, n;\ \varphi_{ik}(w) = \varphi_{ki}(w))$ seien irgendwie vorgegebene differenzierbare Funktionen von w; es sei möglich, aus der Gleichung

$$\det\left(z_{ik} - \varphi_{ik}(w)\right) = 0 \tag{14}$$

die Größe w als Funktion der Elemente von $\mathfrak{Z}$ zu bestimmen. $\nu_1, \ldots, \nu_n$ sei eine nichttriviale Lösung des Systems

$$\sum_{i=1}^{n}\left(z_{ik} - \varphi_{ik}(w)\right)\nu_i = 0.$$

[Für w ist hier und im folgenden immer die aus (14) bestimmte Funktion von $\mathfrak{Z}$ einzusetzen; $\nu_1, \ldots, \nu_n$ hängen also ebenfalls von $\mathfrak{Z}$ ab.] $\varphi'_{ik}(w)$ sei die Ableitung von $\varphi_{ik}(w)$. Der Ausdruck

$$\sum_{i,k}\varphi'_{ik}(w)\,\nu_i\nu_k$$

sei von Null verschieden; es werde für $j = 1, \ldots, n$

$$\mu_j = \mu_j(\mathfrak{Z}) = \frac{\nu_j}{\sqrt{\sum_{i,k}\varphi'_{ik}(w)\,\nu_i\nu_k}}$$

gesetzt. Dann gilt

$$\sum_{i,k}\varphi'_{ik}(w)\,\mu_i\mu_k = 1 \tag{15}$$

und

$$\sum_{i}\left(z_{ik} - \varphi_{ik}(w)\right)\mu_i = 0 \qquad (k = 1, \ldots, n). \tag{16}$$

[5]) Bei Einführung neuer Variabeln $\zeta_{ik} = \varepsilon_{ik}z_{ik}$ vereinfachen sich (12) und die daraus hergeleiteten Gleichungen noch etwas durch den Wegfall der ε-Faktoren. — Für $m > 1$ lassen sich entsprechende Gleichungen aufstellen.

Differentiation der daraus folgenden Gleichung

$$\sum_{i,k} \left(z_{ik} - \varphi_{ik}(w)\right) \mu_i \mu_k = 0$$

liefert

$$\sum_{i,k} \left(dz_{ik} - \varphi'_{ik}(w)\, dw\right) \mu_i \mu_k + 2 \sum_{i,k} \left(z_{ik} - \varphi_{ik}(w)\right) \mu_i\, d\mu_k = 0,$$

also unter Benutzung von (15) und (16)

$$\sum_{i,k} dz_{ik}\, \mu_i \mu_k - dw = 0;$$

das ist (12) in anderer Schreibweise.

Für $m=1$, $n=2$ reduziert sich (13) auf

$$4 \frac{\partial w}{\partial z_{11}} \frac{\partial w}{\partial z_{22}} - \left(\frac{\partial w}{\partial z_{12}}\right)^2 = 0.$$

Im euklidischen (z_{11}, z_{12}, z_{22})-Raum gedeutet sind die Flächen $w = \text{const}$ also Torsen, deren Tangentialebenen parallel zu den Tangentialebenen eines festen Kegels verlaufen. Eine Lösung ist beispielsweise $w = 2z_{12} + 2\sqrt{z_{11} z_{22}}$; für dieses w ist

$$dw = \begin{pmatrix} dz_{11} & dz_{12} \\ dz_{12} & dz_{22} \end{pmatrix} \begin{bmatrix} \sqrt[4]{\frac{z_{22}}{z_{11}}} \\ \sqrt[4]{\frac{z_{11}}{z_{22}}} \end{bmatrix}.$$

§ 4. Die Eigenwerte bei rangerniedrigender linearer Abbildung

Jeder m-reihigen komplexen quadratischen Matrix $\mathfrak{M} = \mathfrak{M}^{(1)} + i\,\mathfrak{M}^{(2)}$ werde die $2m$-reihige reelle quadratische Matrix

$$\widetilde{\mathfrak{M}} = \begin{pmatrix} \mathfrak{M}^{(1)} & \mathfrak{M}^{(2)} \\ \mathfrak{M}^{(2)} & -\mathfrak{M}^{(1)} \end{pmatrix}$$

zugeordnet und umgekehrt. Ist $\mathfrak{Z}' = \mathfrak{Z}$, so ist offenbar $\widetilde{\mathfrak{Z}}' = \widetilde{\mathfrak{Z}}$ und umgekehrt. Ist $\mathfrak{U}$ unitär, so ist $\widetilde{\mathfrak{U}}$ orthogonal und umgekehrt, wenn man die Unitaritätsbedingung $\mathfrak{U}'\overline{\mathfrak{U}} = \mathfrak{E}$ (m-reihige Einheitsmatrix) in Real- und Imaginärteil spaltet und dann mit der Orthogonalitätsbedingung $\widetilde{\mathfrak{U}}'\widetilde{\mathfrak{U}} = \mathfrak{E}$ ($2m$-reihige Einheitsmatrix) vergleicht.

Hilfssatz. *Sei* $\mathfrak{W}' = \mathfrak{W}$, *und* $\mathfrak{W}\overline{\mathfrak{W}}$ *habe die (eo ipso reellen und nichtnegativen) Eigenwerte* $q_1^2 \geqq \cdots \geqq q_m^2$; *dabei sei* $q_\mu \geqq 0$ *für* $\mu = 1, \ldots, m$. *Dann hat* $\widetilde{\mathfrak{W}}$ *die Eigenwerte* $q_1 \geqq \cdots \geqq q_m \geqq -q_m \geqq \cdots \geqq -q_1$, *und* $\widetilde{\mathfrak{W}}^2$ *hat die Eigenwerte* $q_1^2 = q_1^2 \geqq q_2^2 = q_2^2 \geqq \cdots \geqq q_m^2 = q_m^2$.

Beweis. Man setze

$$\mathfrak{Q} = \begin{pmatrix} q_1 & & \\ & \ddots & \\ & & q_m \end{pmatrix}.$$

Nach SIEGEL [2], p. 14–15 gibt es eine unitäre Matrix $\mathfrak{U}$ so, daß $\mathfrak{W}=\mathfrak{Q}[\mathfrak{U}]$ ist. Eine einfache Rechnung zeigt dann $\widetilde{\mathfrak{W}}=\widetilde{\mathfrak{Q}}[\widetilde{\mathfrak{U}}]$. Dabei ist $\widetilde{\mathfrak{U}}$ orthogonal; es folgt $\widetilde{\mathfrak{W}}^2=\widetilde{\mathfrak{U}}'\widetilde{\mathfrak{Q}}\,\widetilde{\mathfrak{U}}\,\widetilde{\mathfrak{U}}'\widetilde{\mathfrak{Q}}\,\widetilde{\mathfrak{U}}=\widetilde{\mathfrak{Q}}^2[\widetilde{\mathfrak{U}}]$. Wegen

$$\widetilde{\mathfrak{Q}}=\begin{pmatrix}\mathfrak{Q} & \\ & -\mathfrak{Q}\end{pmatrix} \quad \text{und} \quad \widetilde{\mathfrak{Q}}^2=\begin{pmatrix}\mathfrak{Q}^2 & \\ & \mathfrak{Q}^2\end{pmatrix}$$

folgen die Behauptungen.

Korollar 1. *Ist* $\mathfrak{E}-\mathfrak{W}\overline{\mathfrak{W}}>0$, *so ist* $\mathfrak{E}-\widetilde{\mathfrak{W}}^2>0$.

Korollar 2. *Ist* $\mathfrak{W}'=\mathfrak{W}$ *und Rang* $\mathfrak{W}\leqq 1$, *so ist Rang* $\widetilde{\mathfrak{W}}\leqq 2$.

Hauptsatz. *Die Abbildung* $\mathfrak{Z}\to\mathfrak{W}$ *sei linear homogen und rangerniedrigend, und für* $\mathfrak{E}-\mathfrak{Z}\overline{\mathfrak{Z}}>0$ *sei* $\mathfrak{E}-\mathfrak{W}\overline{\mathfrak{W}}>0$. *Die Eigenwerte von* $\mathfrak{Z}\overline{\mathfrak{Z}}$ *seien* $p_1^2\geqq\cdots\geqq p_n^2$, *die Eigenwerte von* $\mathfrak{W}\overline{\mathfrak{W}}$ *seien* $q_1^2\geqq\cdots\geqq q_m^2$. *Dann ist* $q_k^2\leqq p_k^2$ *für* $1\leqq k\leqq\min(m,n)$ *und (falls* $n<m$ *ist)* $q_k^2=0$ *für* $n<k\leqq m$.

Beweis. Für eine r-reihige reelle symmetrische Matrix $\mathfrak{M}$ lassen sich die nach abnehmender Größe geordneten Eigenwerte bekanntlich ([3], [4]) wie folgt bestimmen: $\mathfrak{x}$ durchlaufe alle Vektoren eines mindestens $(r-k+1)$-dimensionalen Teilraums T_k des r-dimensionalen reellen Vektorraums mit $\mathfrak{x}'\mathfrak{x}=1$. Man bestimme dafür $\operatorname*{Max}_{\mathfrak{x}'\mathfrak{x}=1}\mathfrak{M}[\mathfrak{x}]$; dies Maximum hängt noch von T_k ab. Man variiere nun T_k und bestimme das Minimum dieser Maxima. Die so gefundene Zahl $\operatorname*{Min}_{T_k}\operatorname*{Max}_{\mathfrak{x}'\mathfrak{x}=1}\mathfrak{M}[\mathfrak{x}]$ ist dann der k-te Eigenwert von $\mathfrak{M}$.

Wir wenden dies Verfahren auf die $2m$-reihige Matrix $\widetilde{\mathfrak{W}}^2=\widetilde{\mathfrak{W}}'\widetilde{\mathfrak{W}}$ an. Der $2k$-te Eigenwert dieser Matrix ist nach dem Hilfssatz gleich q_k^2; also ist

$$q_k^2=\operatorname*{Min}_{T_{2k}}\operatorname*{Max}_{\mathfrak{x}'\mathfrak{x}=1}\mathfrak{y}'\mathfrak{y}, \tag{17}$$

wobei $\mathfrak{y}=\widetilde{\mathfrak{W}}\mathfrak{x}$ gesetzt wurde. Nach dem bereits genannten Satz von SIEGEL gilt mit $\mathfrak{P}=\begin{pmatrix}p_1 & & \\ & \ddots & \\ & & p_n\end{pmatrix}$ und einer passenden unitären Matrix $\mathfrak{U}$ die Gleichung $\mathfrak{Z}=\mathfrak{P}[\mathfrak{U}]$. Es werde $\mathfrak{Z}_k=\begin{pmatrix}0 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 0\end{pmatrix}[\mathfrak{U}]$ gesetzt, wobei in der hingeschriebenen Diagonalmatrix das k-te Diagonalelement gleich 1 ist und alle übrigen verschwinden. Das Bild von $\mathfrak{Z}_k$ sei $\mathfrak{W}_k$; wegen der Rangerniedrigung ist Rang $\mathfrak{W}_k\leqq$ Rang $\mathfrak{Z}_k=1$, und wegen der Linearität der Abbildung $\mathfrak{Z}\to\mathfrak{W}$ ist $\mathfrak{W}=p_1\mathfrak{W}_1+\cdots+p_n\mathfrak{W}_n$. Da $p_1,\ldots,p_n$ reell sind, folgt daraus

$$\widetilde{\mathfrak{W}}=p_1\widetilde{\mathfrak{W}}_1+\cdots+p_n\widetilde{\mathfrak{W}}_n.$$

Das obige $\mathfrak{y}$ ist also durch

$$\mathfrak{y}=(p_1\widetilde{\mathfrak{W}}_1+\cdots+p_n\widetilde{\mathfrak{W}}_n)\,\mathfrak{x} \tag{18}$$

bestimmt. Sei $1 \leq k \leq \min(m, n)$. Die linearen Gleichungen

$$\mathfrak{W}_1 \mathfrak{x} = 0, \quad \mathfrak{W}_2 \mathfrak{x} = 0, \ldots, \mathfrak{W}_{k-1} \mathfrak{x} = 0 \tag{19}$$

definieren einen mindestens $(2m - 2k + 2)$-dimensionalen Teilraum $T^{(0)}$ des $2n$-dimensionalen Vektorraums, denn nach Korollar 2 ist Rang $\mathfrak{W}_\nu \leq 2$, so daß also das System (19) höchstens $2(k-1)$ linear unabhängige Gleichungen enthält. Wir verkleinern nun die rechte Seite von (17) nicht, wenn wir an Stelle der Minimumbildung über alle T_{2k} nur den speziellen Raum $T^{(0)}$ heranziehen. Also ist

$$q_k^2 \leq \operatorname{Max} \mathfrak{y}' \mathfrak{y}, \tag{20}$$

wobei $\mathfrak{y}$ durch (18) gegeben ist und $\mathfrak{x}$ alle Vektoren zu durchlaufen hat, die den Bedingungen (19) und der Normierungsbedingung $\mathfrak{x}'\mathfrak{x} = 1$ genügen. Für diese speziellen $\mathfrak{x}$ folgt aus (18) und (19)

$$\mathfrak{y} = (p_k \mathfrak{W}_k + \cdots + p_n \mathfrak{W}_n)\, \mathfrak{x}.$$

Entweder ist nun $p_k = p_{k+1} = \cdots = p_n = 0$; dann ist $\mathfrak{y} = 0$ und gewiß $\mathfrak{y}'\mathfrak{y} \leq p_k^2$. Oder es ist $p_k > 0$; in diesem Falle schreiben wir $\mathfrak{y} = p_k \mathfrak{M} \mathfrak{x}$ mit

$$\mathfrak{M} = \frac{p_k}{p_k} \mathfrak{W}_k + \frac{p_{k+1}}{p_k} \mathfrak{W}_{k+1} + \cdots + \frac{p_n}{p_k} \mathfrak{W}_n.$$

$\mathfrak{M}$ ist dabei das Bild der Matrix

$$\mathfrak{Z}^{(0)} = \frac{p_k}{p_k} \mathfrak{Z}_k + \frac{p_{k+1}}{p_k} \mathfrak{Z}_{k+1} + \cdots + \frac{p_n}{p_k} \mathfrak{Z}_n,$$

und $\mathfrak{Z}^{(0)} \overline{\mathfrak{Z}^{(0)}}$ hat die Eigenwerte $0, \ldots, 0, \left(\frac{p_k}{p_k}\right)^2, \left(\frac{p_{k+1}}{p_k}\right)^2, \ldots, \left(\frac{p_n}{p_k}\right)^2$, die sämtlich zwischen 0 und 1 liegen. Also ist $\mathfrak{E} - \mathfrak{Z}^{(0)} \overline{\mathfrak{Z}^{(0)}} \geq 0$, also $\mathfrak{E} - \mathfrak{M} \overline{\mathfrak{M}} \geq 0$, also nach Korollar 1 $\mathfrak{E} - \mathfrak{M}^2 \geq 0$, also, da $\mathfrak{x}'\mathfrak{x} = 1$ ist, $\mathfrak{x}' \overline{\mathfrak{M}}' \mathfrak{M} \mathfrak{x} = \mathfrak{M}^2[\mathfrak{x}] \leq 1$, also $\mathfrak{y}'\mathfrak{y} \leq p_k^2$. In jedem Fall gilt also $\mathfrak{y}'\mathfrak{y} \leq p_k^2$, und aus (20) folgt die Behauptung $q_k^2 \leq p_k^2$. — Die für den Fall $n < m$ gemachte zusätzliche Behauptung ist trivial, da wegen der vorausgesetzten Rangerniedrigung stets Rang $\mathfrak{W} \leq$ Rang $\mathfrak{Z} \leq n$ ist.

Anmerkung. Ohne die Voraussetzung der Rangerniedrigung ist der vorstehende Satz falsch, wie etwa das Beispiel $\mathfrak{Z} \to \mathfrak{W} = \begin{pmatrix} \mathfrak{Z} & \\ & \mathfrak{Z} \end{pmatrix}$ zeigt.

§ 5. Ein Analogon zum SCHWARZschen Lemma

In diesem Paragraphen sei $\mathfrak{W} = \varphi(\mathfrak{Z})$ eine differentiell rangerniedrigende analytische Matrixfunktion, die den Einheitskreis des $\mathfrak{Z}$-Raumes in den Einheitskreis des $\mathfrak{W}$-Raumes abbildet.

Satz 1 (SIEGEL). *Wird bei $\mathfrak{Z} \to \mathfrak{W}$ der Nullpunkt auf den Nullpunkt abgebildet, so gilt dort $d\mathfrak{W} = f(d\mathfrak{Z})$, wobei f eine lineare homogene Funktion ist, die den Einheitskreis in den Einheitskreis abbildet.*

Beweis. $\mathfrak{W}_1$ sei der lineare Bestandteil der Entwicklung der Elemente von $\mathfrak{W}$ nach Potenzprodukten der Elemente von $\mathfrak{Z}$; also ist $\mathfrak{W}_1 = f(\mathfrak{Z})$ eine lineare homogene Funktion von $\mathfrak{Z}$. An der Stelle $\mathfrak{Z} = 0$ ist dann $d\mathfrak{W} = f(d\mathfrak{Z})$. Es bleibt zu zeigen, daß für $\mathfrak{E} - \mathfrak{Z}\overline{\mathfrak{Z}} > 0$ auch $\mathfrak{E} - \mathfrak{W}_1\overline{\mathfrak{W}}_1 > 0$ ist; für den Beweis dieser Tatsache sei auf SIEGEL [2], p. 11—12 verwiesen. — Bisher wurde die differentielle Rangerniedrigung von φ nicht benötigt; sie liefert uns jedoch unmittelbar den

Zusatz. Das obengenannte f ist eine rangerniedrigende lineare homogene Funktion.

Satz 2. *ds sei ein symplektisches Bogenelement im Einheitskreis des $\mathfrak{Z}$-Raumes, dt sei sein Bild im $\mathfrak{W}$-Raum. Dann ist $dt \leq ds$.*

Beweis. Das Bogenelement ds liege an der Stelle $\mathfrak{Z} = \mathfrak{Z}_0$, und das Bogenelement dt liege an der Stelle $\mathfrak{W} = \mathfrak{W}_0$. Durch passende symplektische linear gebrochene Funktionen, die den Einheitskreis des $\mathfrak{Z}$-Raumes bzw. den Einheitskreis des $\mathfrak{W}$-Raumes auf sich abbilden und alle symplektischen Bogenlängen ungeändert lassen, können $\mathfrak{Z}_0$ und $\mathfrak{W}_0$ in die Nullmatrizen übergeführt werden (vgl. [2]). Diese symplektischen linear gebrochenen Funktionen sind nach § 3 differentiell rangerhaltend und ergeben bei Zusammensetzung mit φ eine differentiell rangerniedrigende Funktion, die den Einheitskreis des $\mathfrak{Z}$-Raumes in den Einheitskreis des $\mathfrak{W}$-Raumes abbildet und den Nullpunkt in den Nullpunkt überführt. Man kann daher ohne Beschränkung der Allgemeinheit $\mathfrak{Z}_0 = 0$ und $\mathfrak{W}_0 = 0$ annehmen. Nach (8) ist im Nullpunkt $ds^2 = 4 \cdot \text{Spur}\,(d\mathfrak{Z}\,d\overline{\mathfrak{Z}})$ und entsprechend $dt^2 = 4 \cdot \text{Spur}\,(d\mathfrak{W}\,d\overline{\mathfrak{W}})$. Die Funktion $\mathfrak{W}_1 = f(\mathfrak{Z})$ erfüllt nach Satz 1 die Voraussetzungen des Hauptsatzes von § 4. Die nach abnehmender Größe geordneten Eigenwerte von $\mathfrak{W}_1\overline{\mathfrak{W}}_1$ sind daher jeweils nicht größer als die entsprechenden nach abnehmender Größe geordneten Eigenwerte von $\mathfrak{Z}\overline{\mathfrak{Z}}$. Wegen Satz 1 gilt dasselbe für den Vergleich der Eigenwerte von $d\mathfrak{W}\,d\overline{\mathfrak{W}}$ und $d\mathfrak{Z}\,d\overline{\mathfrak{Z}}$. Folglich ist $dt^2 \leq ds^2$ wie behauptet.

Aus Satz 2 folgt durch Integration

Satz 3. *Bei der Abbildung $\mathfrak{Z} \to \mathfrak{W} = \varphi(\mathfrak{Z})$ nimmt die symplektische Bogenlänge jeder im Einheitskreis gelegenen Kurve nicht zu.*

Satz 4. *$\mathfrak{Z}^{(1)}$ und $\mathfrak{Z}^{(2)}$ seien zwei Punkte im Einheitskreis; $\mathfrak{W}^{(1)} = \varphi(\mathfrak{Z}^{(1)})$ und $\mathfrak{W}^{(2)} = \varphi(\mathfrak{Z}^{(2)})$ seien ihre Bilder. Dann ist die symplektische Entfernung von $\mathfrak{W}^{(1)}$, $\mathfrak{W}^{(2)}$ höchstens gleich der symplektischen Entfernung von $\mathfrak{Z}^{(1)}$, $\mathfrak{Z}^{(2)}$.*

Beweis. Man verbinde $\mathfrak{Z}^{(1)}$ und $\mathfrak{Z}^{(2)}$ durch einen geodätischen Bogen; dessen symplektische Länge l ist dann gleich der symplektischen Entfernung von $\mathfrak{Z}^{(1)}$, $\mathfrak{Z}^{(2)}$ (vgl. [2]). Bei der Abbildung $\mathfrak{Z} \to \mathfrak{W}$ geht dieser Bogen in eine Kurve über, deren Bogenlänge l' nach Satz 3 höchstens gleich l ist. Die symplektische Entfernung l'' von $\mathfrak{W}^{(1)}$, $\mathfrak{W}^{(2)}$ ist höchstens gleich der Länge l' des letztgenannten Bogens, der ja $\mathfrak{W}^{(1)}$ und $\mathfrak{W}^{(2)}$ verbindet. Also ist $l'' \leq l$ wie behauptet.

Die Frage bleibt offen, ob bzw. unter welchen Bedingungen für die nach abnehmender Größe geordneten Eigenwerte der gemäß (7) gebildeten Matrizen

$$(\mathfrak{Z}^{(1)} - \mathfrak{Z}^{(2)})\,(\mathfrak{E} - \overline{\mathfrak{Z}}^{(2)}\,\mathfrak{Z}^{(1)})^{-1}\,(\overline{\mathfrak{Z}}^{(1)} - \overline{\mathfrak{Z}}^{(2)})\,(\mathfrak{E} - \mathfrak{Z}^{(2)}\,\overline{\mathfrak{Z}}^{(1)})^{-1}$$

und

$$(\mathfrak{W}^{(1)} - \mathfrak{W}^{(2)})\,(\mathfrak{E} - \overline{\mathfrak{W}}^{(2)}\,\mathfrak{W}^{(1)})^{-1}\,(\overline{\mathfrak{W}}^{(1)} - \overline{\mathfrak{W}}^{(2)})\,(\mathfrak{E} - \mathfrak{W}^{(2)}\,\overline{\mathfrak{W}}^{(1)})^{-1}$$

entsprechende Ungleichungen gelten wie im Hauptsatz von § 4.

Literatur

[1] Fuhrmann, A.: Klassen ähnlicher Matrizen als verallgemeinerte Doppelverhältnisse. Math. Z. **62**, 211—240 (1955). — [2] Siegel, C. L.: Symplectic Geometry. Amer. J. Math. **65**, 1—86 (1943). — [3] Fischer, E.: Über quadratische Formen mit reellen Koeffizienten. Mh. Math. u. Phys. **16**, 234—249 (1905). — [4] Courant, R., u. D. Hilbert: Methoden der Mathematischen Physik, Bd. 1, S. 16—18. Berlin 1924.

Göttingen, Mathematisches Institut der Universität

(Eingegangen am 12. September 1955)

WIELANDT, H.
Math. Zeitschr. Bd. 63, S. 478—485 (1956)

Primitive Permutationsgruppen vom Grad $2p$

ISSAI SCHUR zum Gedächtnis

Von

HELMUT WIELANDT

1. *Übersicht.* In dieser Arbeit bedeutet p eine Primzahl und $\mathfrak{G}$ eine primitive Permutationsgruppe des Grades $2p$, welche die Symbole („Ziffern") $1, 2, \ldots, 2p$ vertauscht. $\mathfrak{G}_1$ bezeichne die größte Untergruppe $\mathfrak{G}$, welche die Ziffer 1 fest läßt. Die Anzahl der Transitivitätsgebiete, in welche die $2p$ Ziffern unter der Einwirkung von $\mathfrak{G}_1$ zerfallen, heiße t. Wir beweisen:

(1.1) *Stets ist* $t \leqq 3$.

(1.2) *Ist* $t=3$, *so hat der Grad die Form* $2p=m^2+1$ $(m=1, 3, 5, \ldots)$. *Die drei transitiven Konstituenten von* $\mathfrak{G}_1$ *haben die Grade* 1, $m(m-1)/2$, $m(m+1)/2$. *Als Gruppe von komplexen linearen Transformationen betrachtet enthält* $\mathfrak{G}$ *drei irreduzible Bestandteile mit den Graden* 1, $p-1$, p.

Aus (1.1) und (1.2) folgt unmittelbar das Hauptergebnis:

Jede primitive Gruppe vom Grad $2p \neq m^2+1$ *ist zweifach transitiv.*

Offen bleibt die Frage, ob zu jedem Grade $2p=m^2+1$ primitive, aber nicht zweifach transitive Gruppen existieren. Für den kleinsten in Betracht kommenden Wert $p=5$ gibt es solche Gruppen; man erhält zwei Beispiele, wenn man die symmetrische und die alternierende Gruppe des Grades 5 auf die 10 ungeordneten Paare wirken läßt, die aus 5 Ziffern gebildet werden können.

Der Beweis von (1.1) beruht auf einem Imprimitivitätskriterium (2.2), welches vom Verfasser mit Hilfe der Methode von SCHUR [6] gefunden worden ist, aber hier auf einem anderen, elementareren Weg bewiesen wird; im übrigen werden die Methoden benutzt, mit denen BURNSIDE die zweifache Transitivität der nicht auflösbaren Gruppen vom Grade p nachgewiesen hat [1, S. 341]. Der Satz (1.2) ergibt sich durch Betrachtung der Eigenwerte einer gewissen linearen Substitution, die mit allen Elementen von $\mathfrak{G}$ vertauschbar ist.

2. *Zwei Imprimitivitätskriterien.* Im folgenden bedeutet n eine beliebige natürliche Zahl >1 und Ω die Menge der Ziffern $1, 2, \ldots, n$. Ist $\mathfrak{H}$ eine Gruppe von Permutationen dieser Ziffern und $\alpha \in \Omega$, so bezeichnen wir mit $\mathfrak{H}_\alpha$ die Untergruppe aller $H \in \mathfrak{H}$, die α ungeändert lassen. Das Bild von $\omega \in \Omega$ nach Ausführung von H bezeichnen wir mit ω^H. Entsprechend bezeichnen wir, wenn $\Gamma \subseteq \Omega$ ist, mit $\Gamma^{\mathfrak{H}}$ die Menge aller γ^H $(\gamma \in \Gamma, H \in \mathfrak{H})$; so bedeutet $\alpha^{\mathfrak{H}}$ das Transitivitätsgebiet von $\mathfrak{H}$, welchem α angehört. Die identische Per-

mutation bezeichnen wir mit E. Die Anzahl der Elemente einer beliebigen Menge M schreiben wir $|M|$; z. B. bedeutet $|\mathfrak{H}|$ die Ordnung von $\mathfrak{H}$, und $\mathfrak{H}$ ist genau dann transitiv auf Ω, wenn $|\alpha^{\mathfrak{H}}| = n$ ist.

(2.1) *Sei $\alpha \in \Omega$, $\beta \in \Omega$, $\alpha \neq \beta$ und $\Gamma \subseteq \Omega$, $\Delta \subset \Omega$, $|\Gamma| \geq |\Delta|$. Sei $\mathfrak{H}$ eine auf Ω transitive Permutationsgruppe zusammengesetzter Ordnung. Aus $H \in \mathfrak{H}$ und $\alpha^H \in \Gamma$ folge stets $\beta^H \in \Delta$. Dann ist $\mathfrak{H}$ imprimitiv.*

Beweis. Wir bilden die Menge $\mathfrak{C}$ aller $C \in \mathfrak{H}$, für die $\alpha^C \in \Gamma$ ist, und die Menge $\mathfrak{D}$ aller $D \in \mathfrak{H}$, für die $\beta^D \in \Delta$ ist. Nach Voraussetzung ist $\mathfrak{C} \subseteq \mathfrak{D}$. Ferner gilt $\mathfrak{H}_\alpha \mathfrak{C} = \mathfrak{C}$, $\mathfrak{H}_\beta \mathfrak{D} = \mathfrak{D}$. Hieraus folgt wegen der Transitivität von $\mathfrak{H}$

$$|\mathfrak{C}| = |\mathfrak{H}_\alpha|\,|\Gamma| \geq |\mathfrak{H}_\alpha|\,|\Delta| = |\mathfrak{H}_\beta|\,|\Delta| = |\mathfrak{D}|.$$

Wir haben nun $|\mathfrak{C}| \geq |\mathfrak{D}|$ und $\mathfrak{C} \subseteq \mathfrak{D}$, also $\mathfrak{C} = \mathfrak{D}$. Wenn wir also die Gruppe $\mathfrak{M}$ aller $M \in \mathfrak{H}$ bilden, welche die Eigenschaft $M\mathfrak{D} = \mathfrak{D}$ haben, so ist $\mathfrak{H}_\alpha \subseteq \mathfrak{M}$ und $\mathfrak{H}_\beta \subseteq \mathfrak{M}$. Wegen $\Delta \neq \Omega$ ist $\mathfrak{M} \neq \mathfrak{H}$. Nehmen wir nun entgegen der Behauptung von (2.1) an, $\mathfrak{H}$ sei primitiv. Dann sind $\mathfrak{H}_\alpha$ und $\mathfrak{H}_\beta$ maximal in $\mathfrak{H}$, daher ist $\mathfrak{H}_\alpha = \mathfrak{M} = \mathfrak{H}_\beta$. Hiernach enthält der Normalisator von $\mathfrak{H}_\alpha$ nicht nur $\mathfrak{H}_\alpha$, sondern auch alle Elemente $H \in \mathfrak{H}$ mit $\alpha^H = \beta$; er ist also größer als $\mathfrak{H}_\alpha$ und daher gleich $\mathfrak{H}$. Dann ist aber $\mathfrak{H}_\alpha = E$, also E maximal in $\mathfrak{H}$ entgegen der Voraussetzung, daß $|\mathfrak{H}|$ keine Primzahl ist. Damit ist (2.1) bewiesen. (Setzt man darin $\Gamma = \Delta$, so wird, wie man leicht erkennt, die Voraussetzung über $|\mathfrak{H}|$ entbehrlich, und (2.1) geht in einen bekannten Satz von Rudio [4] über.)

Aus (2.1) gewinnen wir das zu Beginn angekündigte Kriterium:

(2.2) *Sei $\mathfrak{H}$ eine auf Ω transitive Permutationsgruppe zusammengesetzter Ordnung. Es gebe in Ω zwei Ziffern $\alpha \neq \beta$ und in $\mathfrak{H}$ eine Untergruppe $\mathfrak{U}$ derart, daß*

$$\beta^{\mathfrak{H}_\alpha} \subseteq \beta^{\mathfrak{U}} \subset \Omega \quad \text{und} \quad |\alpha^{\mathfrak{U}}| \geq |\beta^{\mathfrak{U}}|$$

gilt. Dann ist $\mathfrak{H}$ imprimitiv.

Beweis. Wir setzen $\alpha^{\mathfrak{U}} = \Gamma$, $\beta^{\mathfrak{U}} = \Delta$. Dann sind die Voraussetzungen von (2.1) erfüllt. Denn aus $\alpha^H \in \Gamma$ folgt $\alpha^H = \alpha^U$ mit einem geeigneten $U \in \mathfrak{U}$; es ist $HU^{-1} \in \mathfrak{H}_\alpha$, also

$$\beta^H = \beta^{HU^{-1}U} = (\beta^{HU^{-1}})^U \in \beta^{\mathfrak{H}_\alpha U} \subseteq \beta^{\mathfrak{U}U} = \beta^{\mathfrak{U}} = \Delta.$$

3. *Beschränkung auf gerade Permutationen.* Wir wenden uns nun der zu Beginn genannten primitiven Gruppe $\mathfrak{G}$ des Grades $n = 2p$ zu. Da die zu beweisenden Sätze (1.1) und (1.2) für $p = 2$ trivial sind, nehmen wir $p \geq 3$ an. Ferner können und wollen wir voraussetzen, daß $t \geq 3$, also $\mathfrak{G}$ nicht zweifach transitiv ist; sonst ist nichts mehr zu beweisen. Unter diesen Voraussetzungen behaupten wir:

(3.1) *Die von allen Elementen P der Ordnung p in $\mathfrak{G}$ erzeugte Untergruppe $\mathfrak{G}^*$ hat die folgenden Eigenschaften:*

(a) *$\mathfrak{G}^*$ enthält nur gerade Permutationen.*

(b) *$\mathfrak{G}^*$ ist auf Ω transitiv.*

(c) $\mathfrak{G}^*$ *ist primitiv.*

(d) *Die Anzahl der Transitivitätsgebiete von* $\mathfrak{G}_1^*$ *ist* $t^* \geqq t \geqq 3$.

Beweis. (a) Für jedes erzeugende Element ist $P = P^{p+1}$. Wegen $2 | p+1$ ist P gerade.

(b) p teilt n und daher auch $|\mathfrak{G}|$, also ist $\mathfrak{G}^* \neq E$; ferner ist $\mathfrak{G}^*$ normal in der primitiven Gruppe $\mathfrak{G}$. Daher ist $\mathfrak{G}^*$ transitiv.

(c) Nehmen wir an, $\mathfrak{G}^*$ sei imprimitiv. Dann gibt es zwischen $\mathfrak{G}_1^*$ und $\mathfrak{G}^*$ eine eigentliche Zwischengruppe: $\mathfrak{G}_1^* \subset \mathfrak{K} \subset \mathfrak{G}^*$. Wegen der Transitivität von $\mathfrak{G}^*$ ist der Index $|\mathfrak{G}^* : \mathfrak{G}_1^*| = 2p$, also ist entweder $|\mathfrak{G}^* : \mathfrak{K}| = 2$ oder $|\mathfrak{K} : \mathfrak{G}_1^*| = 2$. Das erste kann nicht eintreten, denn sonst würde $\mathfrak{K}$ als ein Normalteiler des Index 2 in $\mathfrak{G}^*$ alle Elemente $P^{p+1} = P$ enthalten. Folglich ist $|\mathfrak{K} : \mathfrak{G}_1^*| = 2$, und der Normalisator $\mathfrak{L}$ von $\mathfrak{G}_1^*$ in $\mathfrak{G}$ enthält nicht nur, wie sich von selbst versteht, die Gruppe $\mathfrak{G}_1$, sondern $\mathfrak{L}$ enthält auch die nicht in $\mathfrak{G}_1$ liegende Gruppe $\mathfrak{K}$. Wegen der Maximalität von $\mathfrak{G}_1$ ist also $\mathfrak{L} = \mathfrak{G}$. Nun folgt aus der Transitivität von $\mathfrak{G}$, daß $\mathfrak{G}_1^* = E$, also $\mathfrak{G}^*$ regulär ist. Dann besteht aber jedes Element der Ordnung 2 in $\mathfrak{G}^*$ aus p Transpositionen, ist also ungerade, im Widerspruch zu (a).

(d) ist trivial, da $\mathfrak{G}_1^*$ Untergruppe von $\mathfrak{G}_1$ ist.

Nach (3.1) genügt es, den Beweis der Sätze (1.1) und (1.2) für $\mathfrak{G}^*$ statt $\mathfrak{G}$ zu führen. Zur Vereinfachung lassen wir den Stern fort und machen o.B.d.A. die zusätzliche Voraussetzung:

(3.2) $\mathfrak{G}$ *enthält nur gerade Permutationen und wird von Elementen der Ordnung* p *erzeugt.*

4. *Die* p*-Sylow-Normalisatoren.* Unser Beweis des Satzes (1.1) erfordert ein genaues Studium des Normalisators $\mathfrak{N}$ einer p-Sylow-Gruppe $\mathfrak{P}$ von $\mathfrak{G}$. Wir zeigen zunächst:

(4.1) *Kein Element von* $\mathfrak{G}$ *ist ein Zyklus der Ordnung* p.

Beweis. Sonst müßte $\mathfrak{G}$ als eine primitive Gruppe, die eine primitive Untergruppe kleineren Grades enthält, nach einem Satz von JORDAN [3, S. 384] zweifach transitiv sein.

(4.2) $\mathfrak{P}$ *besitzt genau zwei Transitivitätsgebiete* Γ, Δ. *Es ist* $|\Gamma| = |\Delta| = |\mathfrak{P}| = p$.

Dies folgt aus (4.1) in Verbindung mit $n = 2p$.

(4.3) *Kein Element* A *von* $\mathfrak{G}$ *hat eine Ordnung* pq *mit* $q > 1$.

Beweis. Sonst hat A^q die Ordnung p, besteht also nach (4.2) aus zwei Zyklen der Ordnung p. Daher treten in der Zyklenzerlegung von A nur solche Zyklen auf, deren Ordnungen Vielfache von p (aber nicht alle gleich p) sind. Wegen $n = 2p$ kann A nur aus einem einzigen Zyklus der Ordnung $2p$ bestehen. Dann ist aber A ungerade, im Widerspruch zu (3.2).

(4.4) *Es ist* $\Gamma^{\mathfrak{N}} \subseteqq \Gamma$ *und* $\Delta^{\mathfrak{N}} \subseteqq \Delta$.

Beweis. Sonst gäbe es ein $N \in \mathfrak{N}$, welches Γ mit Δ vertauscht; nach (3.2) wäre N gerade. Sei P ein erzeugendes Element von $\mathfrak{P}$ und $N^{-1} P N = P^r$.

Wir wählen eine gerade Permutation R (nicht notwendig in $\mathfrak{G}$), für welche ebenfalls $R^{-1}PR = P^r$, aber außerdem $\Gamma^R = \Gamma$ gilt. Dann wäre NR^{-1} eine gerade Permutation, welche mit P vertauschbar ist und außerdem Γ mit Δ vertauscht. Eine solche gibt es aber nicht.

(4.5) *Sei* $|\mathfrak{N}| = pf$. *Dann besitzt* $\mathfrak{N}_1$ *außer zwei Transitivitätsgebieten von je einer Ziffer nur noch Transitivitätsgebiete von je* f *Ziffern.*

Beweis. Nach (4.2) und (4.4) ist $|1^{\mathfrak{N}}| = p$, daher ist $|\mathfrak{N}_1| = f$. Sei $1 \in \Gamma$. Kein Element von $\mathfrak{N}$ hat auf Γ den identischen Konstituenten (E ausgenommen); denn es müßte $\mathfrak{P}$ zentralisieren, also aus einem Zyklus der Ordnung p auf Δ bestehen, entgegen (4.1). Hiernach hat der Konstituent von $\mathfrak{N}_1$ auf Γ die Ordnung f und besitzt daher, weil er im Normalisator der von einem p-Zyklus erzeugten Gruppe liegt, außer einem Transitivitätsgebiet der Länge 1 nur solche der Länge f. Das gleiche gilt für den Konstituenten von $\mathfrak{N}_1$ auf Δ.

5. *Abschätzung von* t. Als Vorstufe zu (1.1) beweisen wir:

(5.1) *Es ist* $t < 2 + (p-1)/f$.

Beweis. (a) Ist $\beta \in \Omega$, $\beta \neq 1$, so ist $|\beta^{\mathfrak{G}_1}| > 1$. Das ergibt sich z. B. dadurch, daß man in (2.2) $\mathfrak{H} = \mathfrak{G}$, $\alpha = 1$ und $\mathfrak{U} = \mathfrak{G}_1$ setzt.

(b) Nach (a) besteht jedes von 1 verschiedene Transitivitätsgebiet von $\mathfrak{G}_1$ aus mindestens zwei Ziffern. Daraus folgt $t < p + 1$, womit insbesondere der Fall $f = 1$ erledigt ist. Ist $f > 1$, so ziehen wir Satz (4.5) heran. Er besagt, daß $\mathfrak{N}_1$ außer 1 noch eine weitere Ziffer fest läßt (etwa β) und sonst nur Transitivitätsgebiete aus je f Ziffern besitzt. Nach (4.4) ist jedes Transitivitätsgebiet von $\mathfrak{N}_1$ ganz in einem Transitivitätsgebiet von $\mathfrak{P}$ enthalten. Die Transitivitätsgebiete von $\mathfrak{G}_1$ seien $T_1 = 1$, $T_2 \ni \beta$, $T_3, \ldots, T_t$. Dann gilt:

(c) Es ist $|T_2| > 1$ nach (a).

(d) Es ist $|T_2| \geq f + 1$. Denn es gibt $\gamma \in T_2$, $\gamma \neq \beta$, und damit haben wir $\gamma^{\mathfrak{N}_1} \subseteq \gamma^{\mathfrak{G}_1} = T_2$. Nach (4.5) ist $|\gamma^{\mathfrak{N}_1}| = f$.

(e) Für $\tau \geq 3$ ist $|T_\tau| \geq 2f$. Denn wegen $\mathfrak{N}_1 \subseteq \mathfrak{G}_1$ besteht T_τ aus vollen Transitivitätsgebieten von $\mathfrak{N}_1$; von diesen hat jedes die Länge f; T_τ besteht nicht nur aus einem einzigen Transitivitätsgebiet von $\mathfrak{N}_1$, sonst wäre $\mathfrak{G}$ wieder imprimitiv [in (2.2) wähle man $\alpha = 1$, $\beta \in T_\tau$, $\mathfrak{H} = \mathfrak{G}$, $\mathfrak{U} = \mathfrak{P}$].

(f) Aus den Aussagen (d) und (e) folgt

$$1 + (f+1) + (t-2) \cdot 2f \leq \sum_{\tau=1}^{t} |T_\tau| = 2p.$$

Hieraus ergibt sich die Behauptung von Satz (5.1): Es ist

$$t - 2 \leq (2p - 2 - f)/2f < (2p-2)/2f = (p-1)/f.$$

6. *Die irreduziblen Bestandteile von* $\mathfrak{G}$. Wir sehen die Elemente $G \in \mathfrak{G}$ nun nicht mehr als Permutationen auf Ω an, sondern als $2p$-reihige Matrizen, die in jeder Zeile und in jeder Spalte genau einmal 1 und sonst überall 0 haben. Die absolut irreduziblen Bestandteile von $\mathfrak{G}$ nennen wir $\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_r$, ihre Grade $f_1, f_2, \ldots, f_r$ und die Spur der Matrix $\mathfrak{D}_\varrho(G)$ den Charakter $\chi_\varrho(G)$.

Die Numerierung sei so gewählt, daß $\mathfrak{D}_1$ die identische Darstellung ist (daher $f_1 = 1$). Da $\mathfrak{G}$ transitiv ist, ist für $\varrho > 1$ stets $\mathfrak{D}_\varrho \neq \mathfrak{D}_1$, und da $\mathfrak{G}$ nicht zweifach transitiv ist, ist $r > 2$ [1, S. 275]. Mit ζ bezeichnen wir eine fest gewählte primitive p-te Einheitswurzel, mit $\mathfrak{P}$ eine bestimmte p-Sylow-Gruppe von $\mathfrak{G}$.

(6.1) *Ist $E \neq P \in \mathfrak{P}$, so hat P, als Matrix betrachtet, die Eigenwerte $1, \zeta, \ldots, \zeta^{p-1}$, jeden genau zweifach.*

Beweis. P zerfällt nach (4.2) in zwei Zyklen der Ordnung p.

Nach (6.1) gibt es genau ein $\varrho > 1$, für welches die Matrix $\mathfrak{D}_\varrho(P)$ den Eigenwert 1 besitzt. Die Numerierung sei so gewählt, daß $\varrho = 2$ ist. Wir behaupten:

(6.2) *$\chi_2(G)$ ist rational für jedes $G \in \mathfrak{G}$.*

Beweis. Da die Spuren von $\mathfrak{G}$ rational sind, enthält $\mathfrak{G}$ neben jedem Charakter auch alle konjugierten Charaktere. Beim Übergang zu den algebraisch konjugierten Darstellungen muß $\mathfrak{D}_2$ bis auf Äquivalenz fest bleiben, da $\mathfrak{D}_2$ die einzige in $\mathfrak{G}$ enthaltene irreduzible Darstellung $\neq \mathfrak{D}_1$ ist, in welcher P den rationalen Eigenwert 1 hat.

(6.3) *Es ist $f_2 = p$. Die Eigenwerte von $\mathfrak{D}_2(P)$ sind $1, \zeta, \ldots, \zeta^{p-1}$.*

Beweis. Sei $k \geqq 0$ die Vielfachheit, mit der ζ als Eigenwert von $\mathfrak{D}_2(P)$ auftritt. Da $\mathfrak{D}_2$ zu den algebraisch konjugierten Darstellungen äquivalent ist, besitzt $\mathfrak{D}_2(P)$ jeden Wert $\zeta^v \neq 1$ als Eigenwert der Vielfachheit k. Es ist also $f_2 = 1 + k(p-1)$. Wegen $r > 2$ haben wir $f_2 < 2p - 1$, also $k < 2$. Andererseits ist $k > 0$; sonst würden $\mathfrak{D}_2$ und $\mathfrak{D}_1$ auf $\mathfrak{P}$ übereinstimmen, also auch auf den zu $\mathfrak{P}$ konjugierten Untergruppen; da diese nach (3.2) die ganze Gruppe $\mathfrak{G}$ erzeugen, wäre $\mathfrak{D}_2 = \mathfrak{D}_1$. Es bleibt nur die Möglichkeit $k = 1$, also $f_2 = p$.

(6.4) *Sei $3 \leqq \varrho < \sigma \leqq r$. Dann sind $\mathfrak{D}_\varrho$ und $\mathfrak{D}_\sigma$ nicht äquivalent; aber für jedes Element A in $\mathfrak{G}$, dessen Ordnung $a \neq p$ ist, gilt $\chi_\varrho(A) = \chi_\sigma(A)$. Insbesondere ist $f_\varrho = f_\sigma$.*

Beweis. Wir setzen $\chi_3 + \chi_4 + \cdots + \chi_r = \chi$ und bezeichnen eine zugehörige Darstellung von $\mathfrak{G}$ mit $\mathfrak{D}$. Nach (6.1) sind alle Eigenwerte von $\mathfrak{D}(P)$ voneinander verschieden. Daher sind $\mathfrak{D}_\varrho$ und $\mathfrak{D}_\sigma$ nicht äquivalent. Sei ζ^k ein Eigenwert von $\mathfrak{D}_\varrho(P)$ und ζ^l einer von $\mathfrak{D}_\sigma(P)$. Nach (4.3) ist $(p, a) = 1$, also besitzt der Körper der $|\mathfrak{G}|$-ten Einheitswurzeln einen Automorphismus ψ mit $\psi[\zeta^k] = \zeta^l$ und $\psi[\chi_\varrho(A)] = \chi_\varrho(A)$. Der Charakter χ enthält den Charakter $\psi[\chi_\varrho]$, da er rational ist und χ_ϱ enthält. In der zu $\psi[\chi_\varrho]$ gehörigen Darstellung besitzt P den Eigenwert ζ^l. Da $\mathfrak{D}_\sigma$ der einzige Bestandteil von $\mathfrak{D}$ ist, in dem P diesen Eigenwert hat, ist $\psi[\chi_\varrho] = \chi_\sigma$. Insbesondere ist $\chi_\varrho(A) = \psi[\chi_\varrho(A)] = \chi_\sigma(A)$, wie behauptet.

(6.5) *Es ist $t = r$.*

Beweis. In jeder transitiven Permutationsgruppe $\mathfrak{H}$ ist die Anzahl der Transitivitätsgebiete von $\mathfrak{H}_1$ gleich der Quadratsumme der Vielfachheiten der

irreduziblen Bestandteile von $\mathfrak{H}$ [1, S. 275]. Da die r irreduziblen Bestandteile von $\mathfrak{G}$ alle voneinander verschieden sind, hat die Quadratsumme in unserem Fall den Wert r.

7. *Beweis des Satzes* (1.1). Nehmen wir an, entgegen der Behauptung von (1.1) sei $t>3$. Dann ist nach (6.5) auch $r>3$, und wir können (6.4) auf $\varrho=3$, $\sigma=4$ anwenden. Die Orthogonalitätsrelationen ergeben

$$g=|\mathfrak{G}|=\sum_{G\in\mathfrak{G}}\chi_3(G)\{\bar{\chi}_3(G)-\bar{\chi}_4(G)\}=\sum_{\mathfrak{Q}}\sum_{Q\in\mathfrak{Q}}\chi_3(Q)\{\bar{\chi}_3(Q)-\bar{\chi}_4(Q)\}.$$

Hierbei durchläuft $\mathfrak{Q}$ alle p-Sylow-Gruppen von $\mathfrak{G}$. Das sind die g/pf zu $\mathfrak{P}$ konjugierten Gruppen; daher liefert jede der inneren Summen den gleichen Beitrag und wir erhalten

$$g=\frac{g}{pf}\sum_{P\in\mathfrak{P}}\chi_3(P)\{\bar{\chi}_3(P)-\bar{\chi}_4(P)\}=\frac{g}{pf}\sum_{P\in\mathfrak{P}}|\chi_3(P)|^2=\frac{g}{pf}\cdot pf_3.$$

Hieraus folgt $f_3=f$, also nach (6.4) weiter $f_\varrho=f$ $(3\leq\varrho\leq r)$. Nun ergibt sich für den Grad nach (6.3) und (6.5) die Zerlegung

$$2p=1+p+(r-2)f=1+p+(t-2)f.$$

Hieraus folgt $t=2+(p-1)/f$, im Widerspruch zu (5.1). Damit ist (1.1) bewiesen.

8. *Beweis des Satzes* (1.2). Von nun an sei $t=3$, also auch $r=3$.

(8.1) *Die beiden von* 1 *verschiedenen Transitivitätsgebiete von* $\mathfrak{G}_1$ *seien* Ξ, H. *Es sei* $|\Xi|=x$, $|H|=y$ *und etwa* $x\leq y$. *Dann ist* (a) $2p=1+x+y$ *und* (b) $2<x<p$.

Beweis. Die Behauptung (a) folgt aus $\Omega=\{1\}\cup\Xi\cup H$.

(b) Nach dem Beweisteil (a) von Satz (5.1) ist $x>1$. Nehmen wir $x=2$ an, also etwa $\Xi=\{\alpha,\beta\}$. Dann hat die Gruppe $\mathfrak{G}_{1\alpha}$ in $\mathfrak{G}_1$ den Index 2, ebenso in $\mathfrak{G}_\alpha$; sie ist also normal in beiden Gruppen und daher auch in deren Erzeugnis $\mathfrak{G}$; hieraus folgt $\mathfrak{G}_{1\alpha}=E$, also $|\mathfrak{G}_1|=2$, $y=2$, $2p=5$. Dieser Widerspruch zeigt, daß $x>2$ ist. Andererseits ist $x<p$; sonst wäre $y\geq x\geq p$ und nach (a) $2p\geq 1+p+p$.

(8.2) *Es gibt eine quadratische Matrix* X, *welche in jeder ihrer* $2p$ *Spalten an genau* x *Stellen mit* 1, *sonst überall mit* 0 *besetzt ist und mit jeder Matrix aus* $\mathfrak{G}$ *vertauschbar ist. Es ist* $\mathrm{Spur}\,X=0$ *und* $\mathrm{Spur}\,\bar{X}'X=2px$.

Beweis. Man konstruiert X nach Burnside [2, S. 174] und Schur [6, S. 599], indem man in der ersten Spalte diejenigen Stellen mit 1 besetzt, deren Zeilenindizes in Ξ liegen. Die übrigen Spalten von X sind dann durch die Forderung der Vertauschbarkeit mit allen $G\in\mathfrak{G}$ festgelegt. Dabei erhalten alle Diagonalelemente von X den gleichen Wert, und zwar den Wert 0 wegen $1\notin\Xi$. Daher ist $\mathrm{Spur}\,X=0$. Die Spur von $\bar{X}'X$ ist die Absolutquadratsumme aller Elemente von X, also gleich $2px$.

(8.3) *Aus der in* (8.2) *genannten Matrix* X *werde mittels einer unitären Matrix* U, *für welche* $U\mathfrak{G}U^{-1}$ *vollständig reduziert ist, die Matrix* $Z = UXU^{-1}$ *gebildet. Dann ist* Z *eine Diagonalmatrix, in welcher ein Diagonalelement* z_1 *einfach, ein weiteres Diagonalelement* z_2 p*-fach und ein drittes Diagonalelement* z_3 $(p-1)$*-fach auftritt.*

Beweis. Die Gruppe $U\mathfrak{G}U^{-1}$ ist vollständig reduziert, ihre irreduziblen Bestandteile $\mathfrak{D}_\varrho$ $(\varrho = 1, 2, 3)$ sind paarweise nicht äquivalent und haben die Grade $1, p, p-1$. Jede Matrix dieser Gruppe ist mit Z vertauschbar. Nach dem bekannten Lemma von SCHUR [5, S. 409] hat Z die angegebene Gestalt.

(8.4) *Die Zahlen* z_1, z_2, z_3 *sind ganz rational. Es gilt*

(a) $z_1 + p z_2 + (p-1) z_3 = 0$

(b) $|z_1|^2 + p|z_2|^2 + (p-1)|z_3|^2 = 2px$

(c) $z_2 \neq x, \quad z_3 \neq x$

(d) $z_1 = x$.

Beweis. Die linke Seite von (a) hat den Wert $\mathrm{Spur}\, Z = \mathrm{Spur}\, X = 0$. Die linke Seite von (b) ist $\mathrm{Spur}\, \overline{Z}'Z = \mathrm{Spur}\, \overline{X}'X = 2px$. Wäre $z_2 = x$, so würde aus (b) folgen, daß $px^2 \leqq 2px$ ist, also $x \leqq 2$, entgegen (8.1). Wäre $z_3 = x$, so würden wegen (a) jedenfalls z_1 und z_2 nicht beide verschwinden, also ergäbe sich aus (b) die mit (8.1) unverträgliche Ungleichung $(p-1)x^2 < 2px$. Mindestens eine der Zahlen z_i muß aber mit x übereinstimmen. Denn x ist ein Eigenwert von X, da für die Zeile $\mathfrak{e} = (1, 1, \ldots, 1)$ nach Konstruktion von X die Gleichung $\mathfrak{e}X = x\mathfrak{e}$ gilt. Nach (c) ist $z_1 = x$. Insbesondere ist z_1 ganz rational.

Daß auch z_2 und z_3 ganz rational sind, ergibt sich so: Die charakteristische Gleichung von X hat ganze rationale Koeffizienten. Daher sind die Wurzeln z_2 und z_3 zunächst einmal ganz algebraisch. Ferner ist die genaue Vielfachheit der Wurzel z_3 entweder $p-1$ oder, wenn $z_2 = z_3$ sein sollte, $2p-1$; in jedem Fall ist sie verschieden von der Vielfachheit jeder anderen Wurzel der charakteristischen Gleichung. Folglich ist z_3 rational. Daß auch z_2 rational ist, folgt nun aus (a).

Wir kommen zum Beweis des Satzes (1.2). Lesen wir die unter (8.4a) bewiesene Gleichung mod p, so erhalten wir wegen $z_1 = x$ die Kongruenz

(e) $$x - z_3 \equiv 0 \bmod p.$$

Ferner ergibt sich aus (b) die Abschätzung

$$z_3^2 \leqq 2px/(p-1) \leqq px < p^2.$$

Hiernach ist $|z_3| < p$ und wegen $0 < x < p$ weiter

(f) $$-p < x - z_3 < 2p.$$

Durch (c), (e), (f) ist $x - z_3 = p$ eindeutig bestimmt. Nun können wir vermöge (a) und (d) sowohl x wie z_3 durch z_2 ausdrücken. Das Ergebnis ist

(g) $$x = p - z_2 - 1, \qquad z_3 = -z_2 - 1.$$

Führen wir diese Werte in (b) ein, so ergibt sich

$$z_2^2 + (z_2 + 1)^2 = p.$$

Setzen wir $2z_2 + 1 = m$, so folgt aus der letzten Gleichung

$$2p = m^2 + 1.$$

Aus (g) folgt nun

$$x = \frac{m(m-1)}{2}$$

und hieraus wegen $1 + x + y = 2p$ weiter

$$y = \frac{m(m+1)}{2}.$$

Schließlich erkennen wir aus den beiden letzten Gleichungen, daß $m = y - x \geqq 0$ ist, und nach Definition ist $m = 2z_2 + 1$ ungerade; m ist also eine Zahl aus der Reihe $1, 3, 5, \ldots$. Damit sind alle Behauptungen des Satzes (1.2) bewiesen.

Literatur

[1] Burnside, W.: Theory of groups of finite order, 2. Aufl. Cambridge University Press 1911. Dover Publications, New York 1955. — [2] Burnside, W.: On some properties of groups of odd order. Proc. Lond. Math. Soc. **33**, 162—185 (1901). — [3] Jordan, C.: Théorèmes sur les groupes primitifs. J. Math. pures appl. **16**, 383—408 (1871). — [4] Rudio, F.: Über primitive Gruppen. J. reine u. angew. Math. **102**, 1—8 (1888). — [5] Schur, I.: Neue Begründung der Theorie der Gruppencharaktere. Sitzgsber. preuß. Akad. Wiss. **1905**, 406—432. — [6] Schur, I.: Zur Theorie der einfach transitiven Permutationsgruppen. Sitzgsber. preuß. Akad. Wiss., phys.-math. Kl. **1933**, 598—623.

Tübingen, Mathematisches Institut der Universität

(Eingegangen am 6. September 1955)

Rado, R.
Math. Zeitschr. Bd. 63, S. 486—495 (1956)

Minimal Points of Convex Sets in Sequence Spaces

In memoriam Issai Schur

By

Richard Rado

1. By means of applications of the Hahn-Banach theorem W. W. Rogosinski has obtained new results in the theory of linear equations in infinitely many variables and, in particular, the following theorem ([1], Theorem B).

Theorem 1. *Let m be a positive integer and, for $0 \leq r < m$; $\lambda \geq 0$, $a_{r\lambda}$ and a_r be complex numbers such that the sequences $(a_{r0}, a_{r1}, \ldots)$ are linearly independent and*

$$\sum_{\lambda=0}^{\infty} |a_{r\lambda}| < \infty \qquad (r < m). \tag{1}$$

Let X be the set of all bounded sequences $x = (x_0, x_1, \ldots)$ of complex numbers such that

$$\sum_{\lambda} a_{r\lambda} x_\lambda + a_r = 0 \qquad (r < m). \tag{2}$$

Then there exists an index λ_0 and a number ξ such that, for every x of X,

$$\sup_{\lambda} |x_\lambda| \geq |\xi|, \tag{3}$$

with equality for at least one x of X; also, whenever there is equality in (3) for some x of X, then $x_{\lambda_0} = \xi$.

After being told of this result the present author noticed that it could also be deduced by means of the methods of classical analysis. This note is concerned with more general results obtained by developing such methods further. Rogosinski's methods yield generalisations of Theorem 1 in directions different from those considered here.

2. Before formulating the first generalization of Theorem 1 I introduce some definitions. The set of all sequences of complex numbers is denoted by S. The members of the sequence x will always be denoted by $x_0, x_1, \ldots$, and similarly in the case of y, z, etc. If N is a non-empty set of non-negative integers then I put[1])

$$\|x\|_N = \sup\,[\lambda \in N]\, |x_\lambda|,$$

and if N is empty, i.e. $N = 0$, then I put $\|x\|_N = 0$. Also, $\|x\|_N = \|x\|$, if $N = \{0, 1, \ldots\}$. The statement

$$x = y \quad in \ M$$

means that $x_\lambda = y_\lambda$ for $\lambda \in M$.

[1]) I write $\sup\,[\lambda \in N]\, |x_\lambda|$ instead of $\sup_{\lambda \in N} |x_\lambda|$, and similarly in other cases.

THEOREM 2. *Let m be a positive integer and let a_r be complex numbers and A a $(m\times\infty)$-matrix whose elements $a_{r\lambda}$ $(r<m;\ \lambda=0,1,\ldots)$ are complex numbers satisfying* (1). *Let X be the set of all sequences x of S such that* (2) *holds and the series in* (2) *are absolutely convergent. Suppose that $X\neq 0$. Then there is a sequence z of X, a positive integer n and a partition of the set $\{0,1,\ldots\}$ into disjoint non-empty sets $N_0, N_1, \ldots, N_{n-1}$ such that the following conditions hold.*

$$|z_\lambda| = \varrho_\nu \qquad (\nu<n;\ \lambda\in N_\nu);$$

$$\varrho_0 > \varrho_1 > \cdots > \varrho_{n-1}.$$

If $x\in X-\{z\}$, then there is $\nu<n$ such that $x=z$ in $N_0+\cdots+N_{\nu-1}$,

$$\|x\|_{N_\nu+\cdots+N_{n-1}} > \varrho_\nu.$$

Furthermore, $1\leqq n\leq 1+RkA$, where RkA denotes the rank of the matrix A. There is exactly one sequence z of X corresponding to which, with any value of n, sets N_ν and numbers ϱ_ν can be found having the properties stated.

It can be seen that if $RkA=m$ then, for any choice of numbers a_r, $X\neq 0$.

In order to deduce Theorem 1 from Theorem 2 choose any element λ_0 of N_0 and put $\xi=z_{\lambda_0}$. Then the assertion of Theorem 1 holds. For, let $x\in X-\{z\}$ and let ν have the property stated in Theorem 2. If $\nu=0$, then $\|x\|>\varrho_0=|\xi|=\|z\|$. If $\nu>0$ then $x=z$ in N_0; $x_{\lambda_0}=z_{\lambda_0}=\xi$. This proves Theorem 1.

3. I introduce some further definitions. ω_0 and ω_1 denote the first denumerable and the first non-denumerable ordinal number respectively. For any ordinals k, l I denote by $[k, l]$ the set of all ordinals λ in the range $k\leqq\lambda<l$. Thus

$$[k,l] = \{\lambda : k\leqq\lambda<l\}.$$

Corresponding to every sequence x I define a, possibly transfinite, sequence of successive maxima ϱ_ν of x and some related concepts.

Consider a fixed x of S. Denote by R the set of all non-negative integers λ_0 which have the property that, for suitable non-negative integers $\lambda_1, \lambda_2, \ldots$,

$$|x_{\lambda_0}| < |x_{\lambda_1}| < \cdots,$$

and by L the set of all non-negative integers outside R. In symbols:

$$R = \{\lambda_0 : |x_{\lambda_0}| < |x_{\lambda_1}| < \cdots\}; \qquad L = [0,\omega_0] - R.$$

Then the set $\{-|x_\lambda| : \lambda\in L\}$ is a well-ordered initial segment of the set $Q=\{-|x_\lambda| : \lambda<\omega_0\}$, ordered by magnitude, and the corresponding terminal segment $\{-|x_\lambda| : \lambda\in R\}$ of Q has no first element. Thus L is the maximal well-ordered initial segment of the ordered set Q. Now I can write

$$\{|x_\lambda| : \lambda\in L\} = \{\varrho_\nu : \nu<n\},$$

where n is some ordinal number, finite or infinite, such that

$$n<\omega_1; \qquad \varrho_\mu > \varrho_\nu \quad (\mu<\nu<n).$$

I put

$$N_\nu = \{\lambda : |x_\lambda| = \varrho_\nu\}; \qquad L_\nu = \sum [\mu < \nu] N_\mu;$$

$$R_\nu = [0, \omega_0] - L_\nu \qquad (\nu < n); \qquad \varrho = ||x||_R.$$

Then $N_\nu \neq 0$; $|x_\lambda| = ||x||_{R_\nu} = \varrho_\nu$ $(\nu < n;\ \lambda \in N_\nu)$;

$$|x_\lambda| < \varrho \qquad (\lambda \in R).$$

To complete the definitions I put

$$N_\nu = 0; \quad R_\nu = R; \quad L_\nu = L; \quad \varrho_\nu = \varrho \qquad (n \leq \nu < \omega_1).$$

Then

$$\varrho_\mu \geq \varrho_\nu \qquad (\mu < \nu < \omega_1).$$

If I want to express the fact that the numbers n, ϱ_ν, ϱ and the sets N_ν, L_ν, R_ν, L, R refer to the sequence x, I write $n(x)$, $\varrho_\nu(x)$, ...

For sets M, N the relation $M \subset N$ denotes inclusion, in the wide sense, and the symbol MN the intersection of M and N.

4. Let $z \in X \subset S$. Then z is called a *minimal sequence of* X if the following condition holds. Whenever $x \in X$ then either

(i) $x = z$ in $L(z)$, and $||x||_{R(z)} \geq \varrho(z)$,

or

(ii) there is $\nu < n(z)$ such that $x = z$ in $L_\nu(z)$ and

$$||x||_{R_\nu(z)} > \varrho_\nu(z).$$

The set X is called *weakly closed* if, whenever[2])

$$x^s \in X\,(s < \omega_0),$$

$$\lim [s \to \omega_0]\, x_\lambda^s = y_\lambda \qquad (\lambda < \omega_0),$$

then

$$y \in X.$$

The set X is called *convex* if $x', x'' \in X$ implies $x''' \in X$, where

$$x_\lambda''' = \tfrac{1}{2}(x_\lambda' + x_\lambda'') \qquad (\lambda < \omega_0).$$

It follows that, if X is weakly closed and convex, and

$$y^\tau \in X; \quad ||y^\tau|| \leq c < \infty \qquad (\tau < \omega_0),$$

then

$$y^* \in X,$$

where

$$y_\lambda^* = \sum [\tau < \omega_0]\, 2^{-\tau-1} y_\lambda^\tau \qquad (\lambda < \omega_0).$$

[2]) Throughout this note, upper indices do not denote exponentiation.

5. THEOREM 3. *Let $X \subset S$. Suppose that, for some real number c, the set*

$$Y = \{x : x \in X; \quad ||x|| \leqq c\}$$

is non-empty, convex and weakly closed. Then

(i) *X possesses at least one minimal sequence*

(ii) *if z and w are minimal sequences of X, then*

$$\varrho_\nu(z) = \varrho_\nu(w); \quad N_\nu(z) = N_\nu(w) \ (\nu < \omega_1); \quad n(z) = n(w);$$
$$L(z) = L(w); \quad \varrho(z) = \varrho(w); \qquad z = w \text{ in } L(z).$$

It will be seen that in order to deduce Theorem 2 from Theorem 3, it suffices to show that the set X of Theorem 2 satisfies the hypothesis of Theorem 3 and that, if z is a minimal sequence of X, then $R(z) = 0$; $n(z) \leqq 1 + Rk A$. The proofs of these statements will be given in § 8.

6. We begin by proving (ii) of Theorem 3. Let $\nu < \omega_1$, and suppose that, for $\mu < \nu$,

$$N_\mu(z) = N_\mu(w),$$
$$z = w \quad \text{in} \quad N_\mu(z).$$

Then $R_\nu(z) = R_\nu(w)$. Since z is minimal it follows that

$$||w||_{R_\nu(z)} \geqq ||z||_{R_\nu(z)} = ||z||_{R_\nu(w)}.$$

By symmetry, we obtain $||z||_{R_\nu(w)} \geqq ||w||_{R_\nu(z)}$. Hence $||w||_{R_\nu(z)} = ||z||_{R_\nu(z)}$. Hence, since z is minimal, $w = z$ in $N_\nu(z)$. Now, the definition of $N_\nu(w)$ shows that $N_\nu(w) > N_\nu(z)$. On grounds of symmetry, $N_\nu(z) > N_\nu(w)$ follows and so $N_\nu(z) = N_\nu(w)$. We have thus proved, by induction, that, for $\nu < \omega_1$, $N_\nu(z) = N_\nu(w)$, and $z = w$ in $N_\nu(z)$. Hence $n(z) = n(w)$; $L(z) = L(w)$; $R(z) = R(w)$,

$$z = w \quad \text{in} \quad L(z).$$

Since z is minimal, this implies $||w||_{R(z)} \geqq ||z||_{R(z)} = ||z||_{R(w)}$ and, by symmetry, $||z||_{R(w)} \geqq ||w||_{R(z)}$. Hence $\varrho(z) = ||z||_{R(z)} = ||w||_{R(w)} = \varrho(w)$, and (ii) is established.

7. I now proceed to prove (i) of Theorem 3. As a first step, I define values of symbols

$$x^\nu, \ N_\nu, \ L_\nu, \ R_\nu, \ \varrho_\nu \qquad (\nu < \omega_1). \tag{4}$$

I shall ultimately show that the numbers and sets (4) belong to some sequence z, in the sense defined in § 3. Let $\nu < \omega_1$. Suppose that

$$x^\mu \in Y; \quad N_\mu \subset [0, \omega_0] \qquad (\mu < \nu);$$
$$x^{\mu_0} = x^{\mu_1} \quad \text{in} \quad N_{\mu_0} \qquad (\mu_0 < \mu_1 < \nu).$$

Put

$$L_\nu = \sum [\mu < \nu] N_\mu; \qquad R_\nu = [0, \omega_0] - L_\nu;$$
$$Y_\nu = \prod [\mu < \nu] \{x : x \in Y; \quad x = x^\mu \quad \text{in} \quad N_\mu\}.$$

In particular, $L_0=0$; $Y_0=Y$. I now show that $Y_\nu \neq 0$. If $\nu=0$, then $Y_\nu=Y\neq 0$. If $\nu=\tau+1$, then $x^\tau\in Y_\nu$. Now let ν be a limit ordinal. Then there is $y^\nu\in S$ and a subset

$$\{\mu_{\nu 0}, \mu_{\nu 1}, \ldots\} \tag{5}$$

of $[0,\nu]$ such that $\mu_{\nu 0}<\mu_{\nu 1}<\cdots$,

$$\sup [\tau<\omega_0]\, \mu_{\nu\tau}=\nu,$$

$$\lim [\tau\to\omega_0]\, x_\lambda^{\mu_{\nu\tau}}=y_\lambda^\nu \qquad (\lambda<\omega_0). \tag{6}$$

Then $y^\nu\in Y$, since Y is weakly closed. Let $\mu<\nu$; $\lambda\in N_\mu$. Then there is $\sigma<\omega_0$ such that $\mu_{\nu\sigma}\geqq\mu$. Then, for $\sigma\leqq\tau<\omega_0$,

$$x_\lambda^\mu = x_\lambda^{\mu_{\nu\tau}}; \qquad x_\lambda^\mu = y_\lambda^\nu.$$

Hence

$$y^\nu = x^\mu \text{ in } N_\mu; \qquad y^\nu\in Y_\nu.$$

Put

$$\varrho_\nu = \inf [x\in Y_\nu]\, ||x||_{R_\nu};$$

$$Z_\nu = \{x : x\in Y_\nu;\ ||x||_{R_\nu}=\varrho_\nu\}.$$

Then there is $y^{\nu\tau}\in Y_\nu$; $z^\nu\in S$ such that

$$||y^{\nu\tau}||_{R_\nu}<\varrho_\nu+2^{-\tau} \qquad (\tau<\omega_0),$$

$$\lim [\tau\to\omega_0]\, y_\lambda^{\nu\tau}=z_\lambda^\nu \qquad (\lambda<\omega_0).$$

Then $z^\nu\in Y$. If $\mu<\nu$; $\lambda\in N_\mu$, then $y_\lambda^{\nu\tau}=x_\lambda^\mu$; $z_\lambda^\nu=x_\lambda^\mu$. Hence $z^\nu=x^\mu$ in N_μ; $z^\nu\in Y_\nu$.

It follows from the definition of ϱ_ν that $||z^\nu||_{R_\nu}\geqq\varrho_\nu$. On the other hand, if $\lambda\in R_\nu$, then

$$|y_\lambda^{\nu\tau}|\leqq||y^{\nu\tau}||_{R_\nu}<\varrho_\nu-2^{-\tau},$$

$$|z_\lambda^\nu|\leqq\varrho_\nu; \qquad ||z^\nu||_{R_\nu}\leqq\varrho_\nu.$$

Hence

$$||z^\nu||_{R_\nu}=\varrho_\nu; \qquad z^\nu\in Z_\nu; \qquad Z_\nu\neq 0.$$

Put

$$N_\nu=\bigcap [x\in Z_\nu]\{\lambda : \lambda\in R_\nu;\ |x_\lambda|=\varrho_\nu\}.$$

Now I define x^ν. For every $\tau\in R_\nu-N_\nu$ there is $x^{\nu\tau}\in Z_\nu$ such that $|x_\tau^{\nu\tau}|<\varrho_\nu$. For $\tau\in[0,\omega_0]-(R_\nu-N_\nu)$ I put $x^{\nu\tau}=z^\nu$. Then I put

$$x_\lambda^\nu=\sum [\tau<\omega_0]\, 2^{-\tau-1}x_\lambda^{\nu\tau} \qquad (\lambda<\omega_0).$$

This completes the inductive definition of the entities (4).

If $\lambda<\omega_0$, then

$$|x_\lambda^\nu|\leqq\sum [\tau<\omega_0]\, 2^{-\tau-1}|x_\lambda^{\nu\tau}|\leqq\sum [\tau<\omega_0]\, 2^{-\tau-1}c=c.$$

Hence, since Y is convex and weakly closed, $x^\nu\in Y$. If $\mu<\nu$; $\lambda\in N_\mu$, then $x^{\nu\tau}\in Z_\nu\subset Y_\nu$; $x_\lambda^{\nu\tau}=x_\lambda^\mu$,

$$x_\lambda^\nu=\sum [\tau<\omega_0]\, 2^{-\tau-1}x_\lambda^{\nu\tau}=x_\lambda^\mu, \qquad x^\nu=x^\mu \text{ in } N_\mu; \qquad x^\nu\in Y_\nu.$$

Hence, by definition of ϱ_ν, $||x^\nu||_{R_\nu} \geq \varrho_\nu$. If $\lambda \in R_\nu$, then $|x_\lambda^{\nu\tau}| \leq ||x^{\nu\tau}||_{R_\nu} \leq \varrho_\nu$; $|x_\lambda^\nu| \leq \varrho_\nu$,

$$||x^\nu||_{R_\nu} \leq \varrho_\nu.$$

Hence $||x^\nu||_{R_\nu} = \varrho_\nu$; $x^\nu \in Z_\nu$. If $\tau \in R_\nu - N_\nu$ then, by definition of $x^{\nu\tau}$, $|x_\tau^{\nu\tau}| < \varrho_\nu$. Hence

$$|x_\tau^\nu| \leq 2^{-\tau-1}|x_\tau^{\nu\tau}| + \sum [\sigma \neq \tau]\, 2^{-\sigma-1}|x_\tau^{\nu\sigma}| < 2^{-\tau-1}\varrho_\nu + \sum [\sigma \neq \tau]\, 2^{-\sigma-1}\varrho_\nu = \varrho_\nu,$$

$$|x_\tau^\nu| < \varrho_\nu \qquad (\tau \in R_\nu - N_\nu). \tag{7}$$

Now I show, by induction with respect to ν, that

$$\varrho_\mu \geq \varrho_\nu \qquad (\mu < \nu < \omega_1). \tag{8}$$

Assume that, for some fixed $\nu < \omega_1$ and every μ, ν_0 such that $\mu < \nu_0 < \nu$, the relation $\varrho_\mu \geq \varrho_{\nu_0}$ holds. Let $\mu < \nu$. If $\nu = \tau + 1$, then $x^\tau \in Y_\nu$,

$$\varrho_\nu \leq ||x^\tau||_{R_\nu} \leq ||x^\tau||_{R_\tau} = \varrho_\tau \leq \varrho_\mu.$$

Now let ν be a limit number. Then there is $\tau_0 < \omega_0$ such that $\mu_{\nu\tau_0} \geq \mu$, where $\mu_{\nu\tau_0}$ is the number occurring in (5). Then, for $\lambda \in R_\nu$; $\tau_0 \leq \tau < \omega_0$,

$$|x_\lambda^{\mu_{\nu\tau}}| \leq ||x^{\mu_{\nu\tau}}||_{R_\nu} \leq ||x^{\mu_{\nu\tau}}||_{R_{\mu_{\nu\tau}}} = \varrho_{\mu_{\nu\tau}} \leq \varrho_\mu,$$

and $\tau \to \omega_0$ yields $|y_\lambda^\nu| \leq \varrho_\mu$, where y^ν is defined by (6). Hence $\varrho_\nu \leq ||y^\nu||_{R_\nu} \leq \varrho_\mu$, and (8) follows. Clearly,

$$N_\mu N_\nu = 0 \qquad (\mu < \nu < \omega_1).$$

Hence there exists an ordinal $n < \omega_1$ such that

$$n = \min\,[N_\nu = 0]\,\nu.$$

I can now prove that

$$\varrho_\mu > \varrho_\nu \qquad (\mu < \nu < n). \tag{9}$$

Assume, to the contrary, that $\mu < \nu < n$; $\varrho_\mu = \varrho_\nu$. Then

$$\varrho_\mu \geq \varrho_{\mu+1} \geq \varrho_\nu = \varrho_\mu; \qquad \varrho_\mu = \varrho_{\mu+1}.$$

Put $x' = \frac{1}{2}(x^\mu + x^{\mu+1})$. Then $x' \in Z_{\mu+1}$. By definition of n, I can choose a number λ in $N_{\mu+1}$. Then

$$|x_\lambda'| \leq \tfrac{1}{2}|x_\lambda^\mu| + \tfrac{1}{2}|x_\lambda^{\mu+1}| < \tfrac{1}{2}\varrho_\mu + \tfrac{1}{2}\varrho_{\mu+1} = \varrho_{\mu+1}. \tag{10}$$

Here I have used (7) and the fact that $\lambda \in R_\mu - N_\mu$. But (10) contradicts the definition of $N_{\mu+1}$. Hence (9) follows.

I finally show that x^n is a minimal sequence of X. Put $z = x^n$. Then $||z||_{R_n} = \varrho_n \leq \varrho_\nu$ $(\nu < n)$. Also,

$$|z_\lambda| = |x_\lambda^n| < \varrho_n \qquad (\lambda \in R_n),$$

in view of (7) and the fact that $N_n = 0$. Furthermore,

$$\{-|z_\lambda| : \lambda < \omega_0\} = A + B,$$

where

$$A = \{-|z_\lambda| : \lambda \in L_n\} = \{-\varrho_\nu : \nu < n\}; \qquad B = \{-|z_\lambda| : \lambda \in R_n\}.$$

A is well-ordered, and B has no least element. Every element of A is less than every element of B, and

$$N_\nu = \{\lambda : |z_\lambda| = \varrho_\nu\} \qquad (\nu < n).$$

All this shows that, in the notation introduced at the end of § 3, $n = n(z)$; $N_\nu = N_\nu(z)$; $\varrho_\nu = \varrho_\nu(z)$.

Now let $x \in X$. There are the following cases.

Case 1. $x = z$ in L_n.

Case 11. $x \overline{\in} Y$. Then $||x|| > c$; $||x||_{R_n} > c \geqq ||z||_{R_n} = \varrho(z)$.

Case 12. $x \in Y$. Then $x \in Y_n$; $||x||_{R_n} \geqq \varrho_n = \varrho(z)$.

Case 2. There is $\nu < n$; $\lambda \in N_\nu$ such that $x = z$ in L_ν, and $x_\lambda \neq z_\lambda$.

Case 21. $x \overline{\in} Y$. Then $||x||_{R_\nu} > c \geqq ||z||_{R_\nu} = \varrho_\nu(z)$.

Case 22. $x \in Y$.

Case 221. $||x||_{R_\nu} > \varrho_\nu$. Then $||x||_{R_\nu} > \varrho_\nu = \varrho_\nu(z)$.

Case 222. $||x||_{R_\nu} \leqq \varrho_\nu$. Then $x \in Z_\nu$. Hence, if $x' = \frac{1}{2}(x+z)$, then $x' \in Z_\nu$; $|x_\lambda| = |z_\lambda| = \varrho_\nu$; $x_\lambda \neq z_\lambda$, and therefore $|x'_\lambda| = |\frac{1}{2}(x_\lambda + z_\lambda)| < \varrho_\nu$. But this contradicts the definition of N_ν.

This proves that z is a minimal sequence of X and so completes the proof of Theorem 3.

8. I now prove Theorem 2. Suppose that the hypotheses of Theorem 2 are satisfied. Let $RkA = p$.

It is possible, by a standard process, to replace the system of equations defining X by an equivalent system

$$\sum [\lambda < \omega_0]\, b_{r\lambda}\, x_\lambda + b_r = 0 \qquad (r < m)$$

which is such that $\sum [\lambda < \omega_0] |b_{r\lambda}| < \infty$ $(r < m)$,

$$b_{r\lambda} = 0 \qquad (p \leq r < m;\ \lambda < \omega_0).$$

It then follows from $X \neq 0$ that $b_r = 0$ $(p \leq r < m)$. Hence there is $y \in X$ such that $y_\lambda = 0$ $(\lambda_0 \leq \lambda < \omega_0)$, for some $\lambda_0 < \omega_0$. Now, if $||y|| = c$, we have $c < \infty$, and the set $Y = \{x : x \in X;\ ||x|| \leq c\}$ is not empty. Also, if x', $x'' \in Y$, then $\frac{1}{2}(x' + x'') \in Y$, so that Y is convex. Y is weakly closed. For, let $x^s \in Y$ $(s < \omega_0)$,

$$\lim [s \to \omega_0]\, x^s_\lambda = x^*_\lambda \qquad (\lambda < \omega_0).$$

Then $|x^s_\lambda| \leq c$; $|x^*_\lambda| \leq c$ and, by uniform convergence,

$$\sum a_{r\lambda}\, x^*_\lambda = \lim [s \to \omega_0] \sum a_{r\lambda}\, x^s_\lambda = -a_r \qquad (r < m).$$

Theorem 3 can be applied to X and yields a minimal sequence z of X. Let $n(z) = n$; $\varrho_\nu(z) = \varrho_\nu$, etc.

If $M, N \subset [0, \omega_0]$, then $[a_{r\lambda}]_{r \in M;\, \lambda \in N}$ denotes the minor of the matrix $A = [a_{r\lambda}]$ obtained by restricting r to M and λ to N. If M and N are of the

same finite cardinal, then $|a_{r\lambda}|_{r\in M;\,\lambda\in N}$ is the determinant of that minor. The conditions on r and λ may also take forms different from $r\in M;\ \lambda\in N$.

Put

$$p_\nu = R\,k\,[a_{r\lambda}]_{r<m;\,\lambda\in R_\nu}.$$

In particular, $p_\nu = 0$ if $R_\nu = 0$. Then $p_\nu \geqq p_{\nu+1}$ $(\nu<\omega_1)$. Let $\nu<\omega_1$; $p_\nu = p_{\nu+1} = q>0$. I want to deduce a contradiction. I may assume that $[0,q]\subset R_{\nu+1}$ and $|a_{r\lambda}|_{r,\lambda<a}\neq 0$. Then there is $x'\in X$ such that $x'=z$ in L_ν; $x'=0$ in $R_\nu[q,\omega_0]$. Put, for $s<\omega_0$,

$$z^s = (1-2^{-s})\,z + 2^{-s}\,x'.$$

Then $z^s\in X$; $z^s = z$ in L_ν. Hence, by definition of ϱ_ν, $||z^s||_{R_\nu}\geqq\varrho_\nu$. Let $\lambda\in R_\nu[q,\omega_0]$. Then

$$|z^s_\lambda| = (1-2^{-s})\,|z_\lambda| \leqq (1-2^{-s})\,\varrho_\nu.$$

If, first of all, $\varrho_\nu>0$, then it follows that

$$||z^s||_{R_\nu[q,\,\omega_0]} \leqq (1-2^{-s})\,\varrho_\nu < \varrho_\nu$$

and therefore

$$||z^s||_{R_\nu[0,\,q]} \geqq \varrho_\nu.$$

This also holds if $\varrho_\nu = 0$. By making $s\to\omega_0$ we obtain

$$||z||_{R_\nu[0,\,q]} \geqq \varrho_\nu.$$

Hence there is $\lambda\in R_\nu[0,q]$ such that $|z_\lambda|\geqq\varrho_\nu$. Then, by definition of N_ν,

$$\lambda\in N_\nu[0,q]\subset L_{\nu+1}\,R_{\nu+1} = 0$$

which is the desired contradiction. This proves that, for $\nu<\omega_1$, either $p_\nu > p_{\nu+1}$ or $p_\nu = p_{\nu+1} = 0$. Therefore

$$p_0 > p_1 > \cdots > p_k = 0,$$

for some $k\leqq p_0$. Then $a_{r\lambda}=0$ $(r<m;\ \lambda\in R_k)$, and X is defined by the system of equations

$$\sum\,[\lambda\in L_k]\,a_{r\lambda}\,x_\lambda + a_r = 0 \qquad (r<m).$$

Then $z'\in X$, where $z'=z$ in L_k; $z'=0$ in R_k. Then, in the notation used in the proof of Theorem 2, $z'\in Y_k$ and, by definition of ϱ_k, $0 = ||z'||_{R_k}\geqq\varrho_k$. If now, $k+1<n$, then $\varrho_k > \varrho_{k+1}\geqq 0$ which is false. Hence

$$n\leqq k+1\leqq p_0+1 = 1+R\,k\,A.$$

In order to complete the proof of Theorem 2 I need only show that $R_n = 0$.

We have $R_n = R_{n+1}$; $p_n = p_{n+1} = 0$; $a_{r\lambda}=0$ $(r<m;\ \lambda\in R_n)$. Let $z''=z$ in L_n; $z''=0$ in R_n. Then $z''\in Y_n$; $0 = ||z''||_{R_n}\geqq\varrho_n$. If, now, $R_n\neq 0$ then we could choose an index λ in R_n. Then $\lambda\in R_n = R_n - N_n$ and therefore $|z_\lambda|<\varrho_n = 0$ which is a contradiction. Hence $R_n = 0$, and Theorem 2 is proved.

9. I conclude this note with some illustrations and some remarks on further extensions.

(i) Let $0 \neq N_0 \subset [0, \omega_0]$; $N_1 = [0, \omega_0] - N_0 \neq 0$; $\varrho_0 > \varrho_1 = 0$. Then the following set X, defined by one linear equation, has $n = 2$ and the given N_ν, ϱ_ν associated with it in the sense of Theorems 2 and 3. Choose numbers $a_{0\lambda} > 0$ $(\lambda \in N_0)$ such that $\varrho_0 \sum [\lambda \in N_0] a_{0\lambda} = 1$, and put $a_{0\lambda} = 0$ $(\lambda \in N_1)$. Let X be the set of all sequences x such that

$$\sum [\lambda < \omega_0] a_{0\lambda} x_\lambda - 1 = 0,$$

where the series is stipulated be to be absolutely convergent. Let z be defined by $z_\lambda = \varrho_\nu$ $(\nu < 2;\ \lambda \in N_\nu)$. Then $z \in X$; $||z|| = \varrho_0$. If $||x|| < \varrho_0$, then

$$|\sum a_{0\lambda} x_\lambda| < 1; \qquad x \overline{\in} X.$$

If $||x|| = \varrho_0$; $\lambda_0 \in N_0$; $x_{\lambda_0} \neq \varrho_0$, then $x \overline{\in} X$. This proves that z is the unique minimal sequence of X, according to Theorem 2, and that $n(z) = 2$; $N_\nu(z) = N_\nu$; $\varrho_\nu(z) = \varrho_\nu$ $(\nu < 2)$.

(ii) Let $0 < n < \omega_1$; $0 \neq N_\nu \subset [0, \omega_0)$ $(\nu < n)$; $\varrho_\nu > \varrho > 0$ $(\nu < n)$;

$$\varrho_\mu > \varrho_\nu; \quad N_\mu N_\nu = 0 \qquad (\mu < \nu < n).$$

Suppose that the set $[0, \omega_0] - \sum [\nu < n] N_\nu = R$ is infinite. For any given n in the range $0 < n < \omega_1$ it is possible to choose N_ν and ϱ_ν, ϱ such that these conditions hold. I now construct a sequence z such that $\varrho(z) = \varrho$; $n(z) = n$;

$$N_\nu(z) = N_\nu; \quad \varrho_\nu(z) = \varrho_\nu \qquad (\nu < n).$$

Put

$$z_\lambda = \varrho_\nu \qquad (\nu < n;\ \lambda \in N_\nu);$$
$$z_\lambda = \varrho\, \lambda(\lambda + 1)^{-1} \qquad (\lambda \in R).$$

Then z has the required properties.

(iii) An examination of the argument used to prove Theorem 3 shows that all that is required of the set of all complex numbers is that it is a vector space over the field of real numbers possessing a metric under which it is closed and for which every closed sphere is strictly convex. It follows that in Theorem 3 the sequences of complex numbers may be replaced by sequences $(x_0, x_1, \ldots)$ such that, for $\lambda < \omega_0$, x_λ is any element of a given vector space S_λ of the type just described.

(iv) The conclusion of Theorem 3 no longer holds if the sequences x are not of type ω_0 but of a non-denumerable well-order type. This is seen by considering the following example. Let X be the set of all sequences

$$x = (x_0, x_1, \ldots, x_{w_0}, \ldots, x_\lambda, \ldots) \qquad (\lambda < \omega_1),$$

of type ω_1, which have that the property that, for $\lambda < \omega_1$, x_λ is a real number in the range $0 \leqq x_\lambda \leqq 1$, and there exists an ordinal $l = l(x) < \omega_1$ such that $x_\lambda = 1$ $(l \leqq \lambda < \omega_1)$. Then the set X is convex and weakly closed but does not

possess any minimal sequence, i.e. a sequence z with properties analogous to those described in the case of sequences of type ω_0. For, we note that, in our case,

$$\varrho_0 = \inf [x \in X] \, ||x|| = 1 .$$

Let $z \in X$. Then there is $\lambda < \omega_1$ such that $z_\lambda = 1$. If, now, x is such that $x_\lambda = 0$, while $x = z$ in $[0, \omega_1] - \{\lambda\}$, then $x \in X$; $||x|| = \varrho_0$, but $x_\lambda \neq z_\lambda = \varrho_0$.

This shows that z is not a minimal sequence of X.

Obviously, when dealing with ω_1-sequences it is necessary to modify the definition of weak closure of a set by taking into consideration weak limits of sequences x^s $(s < \omega_1)$. In this, stronger, sense, the set just defined would not be weakly closed. But even with this stronger definition of closure there does not appear to be an obvious way of extending Theorem 3 to ω_1-sequences. This is due to the fact that Cantor's diagonal process played a prominent part in the proof of Theorem 3, and there does not seem to exist an equally effective analogue of this process in the case when ω_0-sequences are replaced by ω_1-sequences.

Reference

[1] Rogosinski, W. W.: On finite systems of linear equations with an infinity of unknowns. Math. Z. **63**, 97—108 (1955).

The University, Reading (England)

(Eingegangen am 9. Juli 1955)

Fenchel, W.
Math. Zeitschr. Bd. 63, S. 496—506 (1956)

Über konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten

Issai Schur zum Gedächtnis

Von

Werner Fenchel

Das Problem, das im folgenden diskutiert wird, läßt sich für Funktionen von zwei Variablen x_1 und x_2 folgendermaßen formulieren: In einem konvexen Bereich C der x_1x_2-Ebene sei eine Schar (geschlossener oder offener) konvexer Kurven gegeben, die C einfach und lückenlos überdecken. Gefragt wird, unter welchen Bedingungen eine stetige konvexe Funktion $f(x_1, x_2)$ existiert, deren Niveaukurven diese Kurven sind. Ist die Kurvenschar durch eine Gleichung der Form $\varphi(x_1, x_2) = \text{konst.}$ gegeben, wo $\varphi(x_1, x_2)$ eine in C stetige Funktion ist, so läuft dies auf die Aufstellung notwendiger und hinreichender Bedingungen hinaus, die diese Funktion erfüllen muß, wenn es eine stetige monotone Funktion $F(t)$ einer Variablen derart geben soll, daß die Funktion $f(x_1, x_2) = F\big(\varphi(x_1, x_2)\big)$ konvex ist.

Diese Frage, die auf den ersten Blick trivial erscheinen mag, ist erstmalig von B. de Finetti [3] aufgeworfen und für einen beschränkten abgeschlossenen Bereich C und eine Schar geschlossener Kurven eingehend behandelt worden. An anschaulichen Beispielen zeigt de Finetti, daß die Schar konvexer Kurven keineswegs willkürlich vorgeschrieben werden kann, sondern verschiedenartigen Einschränkungen unterworfen ist, die sowohl ihr lokales als auch ihr globales Verhalten betreffen. Notwendige und hinreichende Bedingungen für die Lösbarkeit der gestellten Aufgabe sind implizit in der Forderung enthalten, daß eines der beiden von de Finetti angegebenen Verfahren zur Konstruktion der gesuchten konvexen Funktion zum Ziele führen soll. In einfacher und durchsichtiger Form dürften sich solche Bedingungen kaum formulieren lassen. Es liegt daher nahe zu versuchen, eine etwas bessere Einsicht in ihre Natur dadurch zu gewinnen, daß man die Frage unter geeigneten Differenzierbarkeitsvoraussetzungen behandelt. In der Arbeit de Finettis wird dieser Fall kurz und auf eine Weise erwähnt, die den Eindruck erwecken könnte, daß dann die Frage eine ganz einfache Antwort zulasse ([3] S. 173—174). Dies ist aber keineswegs der Fall, nicht einmal, wenn man die Kurvenschar analytisch annimmt. Es ist daher vielleicht nicht überflüssig, nochmals auf den Gegenstand zurückzukommen.

Im folgenden wird die Fragestellung verallgemeinert, indem Funktionen von beliebig vielen Variablen und ein nicht notwendig beschränkter konvexer Definitionsbereich in Betracht gezogen werden. Andererseits wird vorausgesetzt, daß die gegebene Funktion $\varphi(x_1, \ldots, x_n)$ und die gesuchte $F(t)$,

für die also $F(\varphi(x_1, \ldots, x_n))$ konvex sein soll, zweimal differenzierbar sind. Es werden notwendige und hinreichende Bedingungen für die Existenz einer solchen Funktion $F(t)$ aufgestellt, die das Verhalten von φ teils in den einzelnen Punkten, teils im großen betreffen. Die lokalen Bedingungen (§ 2) bestehen in Ungleichungen und algebraischen Relationen zwischen den ersten und zweiten partiellen Ableitungen von φ und lassen einfache geometrische Deutungen zu. Im Fall von zwei Variablen ergibt sich beispielsweise, daß an einer Stelle, an der die Krümmung der betreffenden Niveaukurve $\varphi(x_1, x_2) = \text{konst.}$ verschwindet, auch die Gausssche Krümmung der Fläche $z = \varphi(x_1, x_2)$ verschwinden muß. Die globalen Bedingungen (§ 3) sind dagegen auch in dem hier betrachteten Fall weniger durchsichtig, und eine weitere Aufspaltung wäre nicht ohne Interesse. Sie enthalten beispielsweise implizit die einfache notwendige Bedingung, daß alle Niveaumannigfaltigkeiten $\varphi(x_1, \ldots, x_n) = \text{konst.}$ dieselben unendlich fernen Punkte haben müssen (§ 5).

Der Verfasser hat sich mit dem de Finettischen Problem auf Anregung von A. W. Tucker beschäftigt und über die genannten Resultate in einer Vorlesungsreihe über konvexe Mengen und Funktionen [2] berichtet, die er 1951 im Rahmen eines vom Office of Naval Research unterstützten Forschungsprogramms an der Universität Princeton gehalten hat.

1. Vorbemerkungen. Es werden reelle Funktionen betrachtet, die in einem Bereich des n-dimensionalen euklidischen Raumes E^n definiert sind. (Die zu behandelnde Aufgabe hat zwar affininvarianten Charakter, es ist aber vielfach bequem, metrische Begriffe als Hilfsmittel heranzuziehen.) Im E^n wird ein rechtwinkliges Koordinatensystem zugrunde gelegt. Der Punkt mit den Koordinaten $x_1, \ldots, x_n$ wird im allgemeinen kurz mit x bezeichnet werden. Sind $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ zwei Punkte und λ eine reelle Zahl, so werden λx und $x + y$ wie üblich durch

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n), \qquad x + y = (x_1 + y_1, \ldots, x_n + y_n)$$

definiert.

Es sei nun C ein (nicht notwendig beschränkter) konvexer Bereich im E^n, der, um unwesentliche Komplikationen zu vermeiden, stets als offen vorausgesetzt wird. Eine in C definierte Funktion $f(x)$ heißt (nach unten) konvex, wenn für beliebige Punkte x und y aus C und $0 \leq \Theta \leq 1$ die Ungleichung

$$f((1 - \Theta) x + \Theta y) \leq (1 - \Theta) f(x) + \Theta f(y)$$

gilt. Dies ist gleichbedeutend damit, daß die Menge derjenigen Punkte $(x, z) = (x_1, \ldots, x_n, z)$ des $(n+1)$-dimensionalen Raumes E^{n+1}, für welche $x \in C$ und $z \geq f(x)$ gilt, konvex ist. Von den wohlbekannten Eigenschaften der in C nicht konstanten konvexen Funktionen kommen die folgenden zur Anwendung:

a) Eine in C konvexe Funktion $f(x)$ ist dort stetig.

b) Eine in C konvexe Funktion $f(x)$ hat entweder überhaupt keine Extrema oder nur ein absolutes Minimum, das in einer relativ zu C abgeschlossenen konvexen Teilmenge von C angenommen wird.

c) Ist $f(x)$ konvex in C, und ist t eine Zahl aus dem Wertebereich von f, so bilden die Punkte x, in denen $f(x) \leqq t$ gilt, eine relativ zu C abgeschlossene konvexe Teilmenge von C.

d) Existieren die partiellen Ableitungen $f_i = \partial f / \partial x_i$, $i = 1, \ldots, n$, einer in C konvexen Funktion $f(x)$ überall in C, so sind sie stetig, so daß f also insbesondere differenzierbar ist, und sie verschwinden gleichzeitig nur in den eventuellen Minimumspunkten von f.

e) Ist die Funktion $f(x)$ zweimal differenzierbar in C, sind also sie selbst und ihre ersten Ableitungen differenzierbar, so ist sie dann und nur dann konvex, wenn die quadratische Form

$$\sum_{i,j=1}^{n} f_{ij}(x)\, \xi_i \xi_j, \qquad f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j},$$

in den Variablen $\xi_1, \ldots, \xi_n$ für jeden Punkt x aus C positiv semidefinit ist.

2. Lokale notwendige Bedingungen. Wir formulieren nun die zu lösende Aufgabe. Gegeben sei eine im offenen konvexen Bereich C zweimal differenzierbare, nicht konstante Funktion $\varphi(x)$. Es werde

$$\alpha = \inf_{x \in C} \varphi(x), \qquad \beta = \sup_{x \in C} \varphi(x)$$

gesetzt, wo die Werte $-\infty$ und $+\infty$ von α bzw. β zugelassen sind. Der Wertevorrat von $\varphi(x)$ ist dann ein Intervall mit den Endpunkten α und β. Gesucht ist eine in diesem Intervall monoton wachsende, zweimal differenzierbare Funktion $F(t)$, für welche $f(x) = F\big(\varphi(x)\big)$ konvex in C ist.

Wir beginnen mit der Herleitung notwendiger Bedingungen. Es sei also $f(x) = F\big(\varphi(x)\big)$ konvex in C. Mit N_α werde die (eventuell leere) Teilmenge von C bezeichnet, in der $\varphi(x) = \alpha$ ist. Mit den Bezeichnungen

$$F' = \frac{dF}{dt}, \qquad \varphi_i = \frac{\partial \varphi}{\partial x_i}$$

folgt aus

$$f_i = F' \varphi_i, \qquad i = 1, \ldots, n, \tag{1}$$

und den Eigenschaften b) und c) von f, daß die Ableitungen φ_i von φ außerhalb N_α nicht gleichzeitig verschwinden können. Wir haben also

$$k^2 = \varphi_1^2 + \cdots + \varphi_n^2 > 0 \qquad \text{für} \qquad x \notin N_\alpha. \tag{2}$$

Wir formulieren dies als Bedingung

I. *Die Funktion $\varphi(x)$ hat entweder keine stationären Werte oder nur ein absolutes Minimum.*

Speziell folgt, daß φ die obere Grenze β nicht annehmen kann, so daß der Wertebereich $\alpha (\leqq) t < \beta$ geschrieben werden kann, wo die Klammern um das Gleichheitszeichen andeuten, daß es dann und nur dann in Frage kommt, wenn φ ein Minimum hat.

Im Hinblick auf das folgende bemerken wir noch, daß aus d), (1) und $F' \geqq 0$

$$F'(t) > 0 \qquad \text{für} \qquad \alpha < t < \beta \tag{3}$$

folgt.

Durch Differentiation von (1) erhält man mit den obigen analogen Bezeichnungen

$$f_{ij} = F'\,\varphi_{ij} + F''\,\varphi_i\,\varphi_j, \qquad i,j = 1, \ldots, n,$$

und aus e) folgt daher, daß die quadratische Form

$$F' \sum_{i,j=1}^{n} \varphi_{ij}\,\xi_i\,\xi_j + F''\left(\sum_{i=1}^{n} \varphi_i\,\xi_i\right)^2$$

an jeder Stelle x von C positiv semidefinit sein muß. Ist N_α nicht leer, so ist dies wegen $F'(\alpha) \geqq 0$ in den Punkten x von N_α stets von selbst erfüllt; denn da diese Minimumstellen von φ sind, gilt $\varphi_i = 0$, $i = 1, \ldots, n$, und die Form $\sum \varphi_{ij}\,\xi_i\,\xi_j$ ist positiv semidefinit. Wir nehmen daher im folgenden an, daß x nicht zu N_α gehört, daß also $\varphi(x) > \alpha$ ist. Dann muß wegen (3) für jedes solche x die quadratische Form

$$Q_\sigma(\xi,\xi) = \sum_{i,j=1}^{n} \varphi_{ij}\,\xi_i\,\xi_j + \sigma\left(\sum_{i=1}^{n} \varphi_i\,\xi_i\right)^2 \tag{4}$$

für ein gewisses $\sigma = \sigma(x)$, nämlich $\sigma = F''/F'$, positiv semidefinit sein. Als unmittelbare Folge ergibt sich:

II. *Für jedes x aus C ist die Form $\sum \varphi_{ij}\,\xi_i\,\xi_j$ unter der Nebenbedingung $\sum \varphi_i\,\xi_i = 0$ positiv semidefinit.*

Für die Punkte x von N_α ist dies nämlich nach einer soeben gemachten Bemerkung eine Folge von I.

Daß II notwendig ist, könnte man natürlich auch daraus schließen, daß die Menge der Punkte x, in denen $\varphi(x) \leqq t$ ist, für jedes t, $\alpha \leqq t < \beta$, wegen § 1, c) konvex sein muß, da sie ja mit der Menge der Punkte, in denen die konvexe Funktion $F(\varphi(x)) \leqq F(t)$ ist, übereinstimmt. Es sei bemerkt, daß die Konvexität dieser Mengen, wie sehr leicht zu sehen, damit gleichbedeutend ist, daß die Funktion φ „quasikonvex" ist. Dies soll besagen, daß sie auf jeder in C enthaltenen abgeschlossenen Strecke ihr Maximum in einem der Endpunkte annimmt, daß also für alle x und y aus C und $0 \leqq \Theta \leqq 1$

$$\varphi\big((1-\Theta)\,x + \Theta\,y\big) \leqq \operatorname{Max}\left[\varphi(x), \varphi(y)\right].$$

Wir haben nun weiter die Bedingung dafür aufzustellen, daß es zu jedem festen, nicht in N_α enthaltenen x eine Zahl σ derart gibt, daß die Form $Q_\sigma(\xi,\xi)$ für alle ξ_i mit $\sum \varphi_i\,\xi_i \neq 0$ nichtnegativ ist, oder anders ausgedrückt, daß

$$\sup_{\sum \varphi_i \xi_i \neq 0} \left(-\frac{\sum \varphi_{ij}\xi_i\xi_j}{(\sum \varphi_i \xi_i)^2}\right) < \infty \tag{5}$$

ist. Zu diesem Zweck betrachten wir das charakteristische Polynom $\Gamma_{Q_\sigma}(\lambda)$ der Form Q_σ. Mit der üblichen Bedeutung von δ_{ij} und der Festsetzung, daß

die Indizes i und j in den folgenden Determinantenformeln die Werte $1, \ldots, n$ durchlaufen sollen, haben wir

$$(6)\quad \begin{cases} \Gamma_{Q_\sigma}(\lambda) = |\varphi_{ij} - \lambda\,\delta_{ij} + \sigma\,\varphi_i\,\varphi_j| \\ \qquad = \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} + \sigma\,\varphi_i\,\varphi_j & \varphi_i \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ -\sigma\,\varphi_j & 1 \end{vmatrix} \\ \qquad = |\varphi_{ij} - \lambda\,\delta_{ij}| - \sigma \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix}. \end{cases}$$

Wir setzen zur Abkürzung $P(\xi, \xi) = \sum \varphi_{ij}\,\xi_i\,\xi_j$ und bezeichnen mit $P^*(\xi, \xi)$ dieselbe Form, wenn die Variablen ξ_i der Bedingung $\sum \varphi_i\,\xi_i = 0$ unterworfen sind. Die charakteristischen Polynome dieser Formen sind

$$\Gamma_P(\lambda) = |\varphi_{ij} - \lambda\,\delta_{ij}|$$

und

$$\Gamma_{P^*}(\lambda) = -\tfrac{1}{k^2} \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix}.$$

Die Richtigkeit der letzteren Gleichung kann man z.B. so einsehen: Beide Seiten sind invariant bei orthogonalen Transformationen der Koordinaten x_i, da solche Transformationen auf orthogonale Transformationen der Variablen ξ_i hinauslaufen (vgl. z.B. [4] S. 241–242). Wählt man nun das Koordinatensystem so, daß an der betreffenden Stelle $\varphi_1 = \cdots = \varphi_{n-1} = 0$, $\varphi_n = k$ wird, so vereinfacht sich die Nebenbedingung zu $\xi_n = 0$, und die rechte Seite geht tatsächlich in das charakteristische Polynom von $\sum_{i,j=1}^{n-1} \varphi_{ij}\,\xi_i\,\xi_j$ über. Hiernach kann (6) also

$$\Gamma_{Q_\sigma}(\lambda) = \Gamma_P(\lambda) + \sigma\,k^2\,\Gamma_{P^*}(\lambda)$$

geschrieben werden. Bezeichnen wir die ν-ten elementarsymmetrischen Funktionen der Eigenwerte von P, P^* und Q_σ mit S_ν, S_ν^* bzw. $T_\nu^{(\sigma)}$ und setzen $S_0 = S_0^* = T_0^{(\sigma)} = 1$, so haben wir

$$\begin{aligned} \Gamma_P(\lambda) &= S_n - S_{n-1}\,\lambda + \cdots + (-1)^n\,S_0\,\lambda^n, \\ \Gamma_{P^*}(\lambda) &= S_{n-1}^* - S_{n-2}^*\,\lambda + \cdots + (-1)^{n-1}\,S_0^*\,\lambda^{n-1}, \\ \Gamma_{Q_\sigma}(\lambda) &= T_n^{(\sigma)} - T_{n-1}^{(\sigma)}\,\lambda + \cdots + (-1)^n\,T_0^{(\sigma)}\,\lambda^n. \end{aligned}$$

Die Forderung, daß Q_σ positiv semidefinit sein soll, ist also gleichbedeutend mit

$$(7)\qquad T_\nu^{(\sigma)} = S_\nu + \sigma\,k^2\,S_{\nu-1}^* \geq 0, \qquad \nu = 1, \ldots, n,$$

wo das Verschwinden eines $T_\nu^{(\sigma)}$ das aller folgenden nach sich zieht.

Es sei nun $r-1$ der (natürlich im allgemeinen von x abhängige) Rang von P^*. Dann haben wir nach II

$$(8)\qquad S_\nu^* > 0 \quad \text{für} \quad 0 \leq \nu \leq r-1, \qquad S_\nu^* = 0 \quad \text{für} \quad r \leq \nu \leq n-1,$$

so daß aus (7) speziell

$$(9)\qquad S_{r+1} \geq 0$$

folgt. Zwischen den Eigenwerten

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

und

$$\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda_{n-1}^*$$

von P bzw. P^* bestehen infolge ihrer Minimum-Maximum-Eigenschaften bekanntlich die Ungleichungen

$$\lambda_1 \geq \lambda_1^* \geq \lambda_2 \geq \cdots \geq \lambda_\nu \geq \lambda_\nu^* \geq \cdots \geq \lambda_{n-1}^* \geq \lambda_n$$

(vgl. z.B. [1] S. 28). Wegen (8) ist nun

$$\lambda_\nu^* > 0 \quad \text{für} \quad 1 \leq \nu \leq r-1, \qquad \lambda_\nu^* = 0 \quad \text{für} \quad r \leq \nu \leq n-1,$$

folglich

$$\lambda_\nu > 0 \quad \text{für} \quad 1 \leq \nu \leq r-1$$

und, falls $r < n$ ist,

$$\lambda_r \geq 0, \quad \lambda_\nu = 0 \quad \text{für} \quad r+1 \leq \nu \leq n-1, \quad \lambda_n \leq 0,$$

insbesondere also $S_{r+1} = \lambda_1 \ldots \lambda_r \lambda_n \leq 0$. Zusammen mit (9) gibt dies

$$S_{r+1} = 0 \quad \text{falls} \quad r < n;$$

d.h. der Rang von P kann höchstens gleich r sein. Wir formulieren dies als weitere notwendige Bedingung:

III. *Hat im Punkt $x \in C$ die Form $\sum \varphi_{ij}\, \xi_i\, \xi_j$ unter der Nebenbedingung $\sum \varphi_i\, \xi_i = 0$ den Rang $r(x) - 1$, so hat sie ohne diese Bedingung höchstens den Rang $r(x)$.*

Dies wurde zwar nur für die nicht zu N_α gehörigen Punkte von C bewiesen; für $x \in N_\alpha$ ist es aber wegen $\varphi_i(x) = 0$, $i = 1, \ldots, n$, von selbst erfüllt.

Die Notwendigkeit einer Bedingung von der Art III ist auch anschaulich unmittelbar einleuchtend. Für $n = 2$ betrachte man in der $x_1 x_2$-Ebene etwa das Geradenbüschel mit dem Nullpunkt als Scheitel. In einem in der Halbebene $x_1 > 0$ enthaltenen konvexen Gebiet C sind die Büschelgeraden die Niveaukurven der Funktion $\varphi = x_2/x_1$. Die Bedingungen I und II sind offensichtlich erfüllt. Für jede wachsende Funktion $z = F(t)$ ist aber die Fläche $z = F(x_2/x_1)$ eine windschiefe Regelfläche und daher niemals konvex. Daß III auch bei Konvexität im engeren Sinne der Niveaukurven nicht überall erfüllt zu sein braucht, zeigt das Beispiel der Funktion $\varphi = (x_2 - x_1^4)/(x_1 + 1)$ in der Halbebene $x_1 > -1$. Auch hier sind I und II erfüllt, während III für $x_1 = 0$ nicht erfüllt ist. Es beruht dies darauf, daß die Niveaukurven $x_2 = x_1^4 + t(x_1 + 1)$ in ihren Schnittpunkten mit der x_2-Achse verschwindende Krümmung, aber nicht parallele Tangenten haben.

Im allgemeinen Fall läßt III folgende geometrische Deutung zu: Verschwinden in einem Punkt der Hyperfläche $\varphi(x) = \text{konst.}$ des F^n genau $n - r$ ihrer Hauptkrümmungen, so müssen an dieser Stelle auch mindestens $n - r$ der Hauptkrümmungen der Hyperfläche $z = \varphi(x)$ des E^{n+1} verschwinden.

3. Globale notwendige Bedingungen. Die Zahlen σ, für welche $Q_\sigma(\xi, \xi)$ positiv semidefinit ist, müssen nach (7) der Ungleichung

$$\sigma \geqq \sigma_0 = \operatorname*{Max}_{1 \leq \nu \leq r} \left(- \frac{S_\nu}{k^2 S^*_{\nu-1}} \right)$$

genügen. Da nun, wie schon bemerkt, das Verschwinden eines $T_\nu^{(\sigma)}$ das aller folgenden nach sich zieht, muß das Maximum für $\nu = r$ erreicht werden, so daß

$$\sigma_0 = - \frac{S_r}{k^2 S^*_{r-1}}.$$

Wie aus (4) hervorgeht, ist σ_0 zugleich die in (5) auftretende obere Grenze.

Wir ziehen nun die Abhängigkeit vom Punkt x in Betracht. Soll es eine Funktion $F(t)$ mit den verlangten Eigenschaften geben, so muß nach (4)

$$\sigma_0(x) \leqq \frac{F''(\varphi(x))}{F'(\varphi(x))}$$

gelten. Da die rechte Seite auf jeder Niveaumannigfaltigkeit konstant ist, ergibt sich als weitere notwendige Bedingung:

IV. *Für jedes feste t mit $\alpha < t < \beta$ ist*

$$G(t) = \sup_{\varphi(x) = t} \left(- \frac{S_r}{k^2 S^*_{r-1}} \right) < \infty.$$

Schließlich verlangt die Aufgabe, daß eine in $\alpha_{(\leqq)} t < \beta$ zweimal differenzierbare Funktion $F(t)$ mit $F'(t) > 0$ für $\alpha < t < \beta$ existiert, so daß für diese t

$$\frac{F''(t)}{F'(t)} \geqq G(t).$$

Ob sich diese Bedingung wesentlich reduzieren läßt, muß dahingestellt bleiben. Da Stetigkeit der zweiten Ableitungen von φ nicht vorausgesetzt ist und die Niveaumannigfaltigkeiten sich ins Unendliche erstrecken können, wird man im allgemeinen von $G(t)$ wohl kaum mehr als die Meßbarkeit behaupten können. Wir formulieren daher als notwendige Bedingung:

V. *Die Funktion $G(t)$, $\alpha < t < \beta$, hat eine Majorante, die die logarithmische Ableitung einer in $\alpha < t < \beta$ positiven und in $\alpha_{(\leqq)} t < \beta$ differenzierbaren Funktion ist.*

4. Das Hinreichen der Bedingungen. Der Beweis dafür, daß die aufgestellten Bedingungen I bis V auch hinreichend sind, ist sehr einfach. Es sei eine in einem offenen konvexen Bereich C des E^n definierte, zweimal differenzierbare Funktion $\varphi(x)$ gegeben, die I bis V erfüllt. Nach V gibt es eine Funktion $H(t)$, die in $\alpha_{(\leqq)} t < \beta$ differenzierbar und in $\alpha < t < \beta$ positiv ist, derart daß

$$\frac{H'(t)}{H(t)} \geqq G(t), \qquad \alpha < t < \beta.$$

Es sei $F(t)$ ein unbestimmtes Integral von $H(t)$. Dann ist $F(t)$ zweimal differenzierbar und monoton wachsend in $\alpha_{(\leqq)} t < \beta$. Wir haben zu zeigen, daß

die quadratische Form

$$F'(\varphi(x))\sum \varphi_{ij}(x)\,\xi_i\xi_j + F''(\varphi(x))\,(\sum \varphi_i\xi_i)^2 \tag{10}$$

für jedes $x \in C$ positiv semidefinit ist.

Besitzt $\varphi(x)$ gemäß I ein Minimum α, so ist diese Behauptung offenbar für die Punkte von N_α richtig; denn dort verschwinden die φ_i, die Form $\sum \varphi_{ij}\,\xi_i\,\xi_j$ ist positiv semidefinit und $F'(\alpha) \geqq 0$. Wir können daher $\varphi(x) > \alpha$ und folglich $F'(\varphi) > 0$ voraussetzen und statt (10) die Form $Q_\sigma(\xi, \xi)$ mit $\sigma = F''/F'$ betrachten. Nun ist nach der obigen Definition von $F(t)$

$$\sigma = \frac{F''(\varphi)}{F'(\varphi)} \geqq G(\varphi) \geqq \sigma_0 = -\frac{S_r}{k^2\,S^*_{r-1}},$$

also

$$Q_\sigma(\xi, \xi) \geqq Q_{\sigma_0}(\xi, \xi).$$

Es genügt also zu zeigen, daß die letztere Form positiv semidefinit ist. Mit den früheren Bezeichnungen haben wir für diese Form

$$T_\nu^{(\sigma_0)} = S_\nu - \frac{S_r}{S^*_{r-1}}\,S^*_{\nu-1}, \qquad \nu = 1, \ldots, n,$$

also

$$T_\nu^{(\sigma_0)} = 0 \qquad \text{für} \qquad r \leqq \nu \leqq n$$

auf Grund der Definition von r und wegen III. Die Form Q_{σ_0} hat also höchstens den Rang $r-1$. Da sie nun für $\sum \varphi_i\,\xi_i = 0$ mit P, also mit P^* übereinstimmt und nach II die letztere Form $r-1$ positive Eigenwerte hat, hat sie selbst $r-1$ positive Eigenwerte und ist daher nichtnegativ, wie behauptet.

5. Beschränktheits- und Unbeschränktheitsrichtungen der Niveaumannigfaltigkeiten konvexer Funktionen. Wie schon in der Einleitung erwähnt, haben die Niveaumannigfaltigkeiten einer konvexen Funktion alle dieselben unendlich fernen Punkte. Dies muß folglich auch für die gegebene Funktion $\varphi(x)$ gelten und daher eine Folge der aufgestellten notwendigen Bedingungen sein. Es läßt sich aber einfach direkt und ohne Differenzierbarkeitsvoraussetzungen zeigen. Zur Beleuchtung der Natur der globalen Bedingungen möge der Beweis dafür und für die damit eng zusammenhängende Eigenschaft der Niveaumannigfaltigkeiten, in denselben Richtungen beschränkt zu sein, hier kurz angegeben werden.

Vorbereitend stellen wir einige, unter teilweise engeren Voraussetzungen im wesentlichen bekannte Sätze über unbeschränkte konvexe Mengen mit kurzen Beweisen zusammen (vgl. hierzu [5], [6], [2]).

Es sei zunächst M eine beliebige nicht beschränkte Punktmenge und $u = (u_1, \ldots, u_n) \neq (0, \ldots, 0)$ ein Vektor im E^n. Wir sagen, M sei beschränkt in der Richtung u, oder u sei Beschränktheitsrichtung von M, wenn es eine Hyperebene $\sum u_i x_i = c$ mit der positiven Normalenrichtung u gibt, derart daß M in dem durch sie bestimmten abgeschlossenen negativen Halbraum enthalten ist, wenn also

$$\sum u_i\,x_i \leqq c \qquad \text{für} \qquad x \in M.$$

Die Hyperebene heiße dann eine Schranke von M; sie kann speziell Stützhyperebene von M sein. Die sämtlichen Beschränktheitsrichtungen u von M machen, etwa vom Nullpunkt aus abgetragen, einen konvexen Kegel $B(M)$ aus, der speziell leer sein kann. Sind nämlich u und v Beschränktheitsrichtungen, so auch $\lambda u+\mu v$ für alle nichtnegativen λ und μ, für welche $\lambda u+\mu v \neq 0$ ist.

Sei nun $x^{(i)}$, $i=1, 2, \ldots$, eine Folge zu M gehöriger Punkte, die sich im Endlichen nicht häufen. Die Folge der Halbstrahlen $\vec{p x}^{(i)}$, die einen willkürlichen festen Punkt p mit den Punkten $x^{(i)}$ verbinden, häuft sich gegen wenigstens einen Halbstrahl. Die Menge aller solchen Häufungsstrahlen ist ein Kegel $A_p(M)$ mit dem Scheitel p. Er gibt die Richtungen an, in denen sich M ins Unendliche erstreckt, und heiße der Asymptotenkegel von M mit dem Scheitel p. Man sieht sehr leicht, daß $A_p(M)$ abgeschlossen ist und daß zwei zu verschiedenen Punkten p gehörige Kegel $A_p(M)$ durch Translation auseinander hervorgehen. Wir können daher kurz von dem Asymptotenkegel von M sprechen.

Zwischen den beiden eingeführten Kegeln besteht die Relation

$$\overline{B(M)} \subseteq B\big(A(M)\big), \tag{11}$$

wo die Überstreichung den Übergang zur abgeschlossenen Hülle bedeutet; denn eine Schranke von M ist, wie unmittelbar aus der Definition von $A_p(M)$ folgt, auch Schranke des Asymptotenkegels mit einem in ihr gelegenen Scheitel, und der Kegel B der Beschränktheitsrichtungen eines Kegels ist abgeschlossen.

Setzen wir nun M als konvex und (der Übersichtlichkeit halber) als abgeschlossen voraus, so ergeben sich wesentlich schärfere Aussagen. Zunächst folgt, daß dann auch $A_p(M)$ konvex ist. Wird ferner p beliebig in M gewählt, so besteht $A_p(M)$ genau aus den von p ausgehenden Halbstrahlen, die ganz zu M gehören. Daß diese Halbstrahlen in $A_p(M)$ enthalten sind, ist nämlich klar. Ist andererseits $\vec{p x}^{(i)}$ eine Folge von Halbstrahlen der oben beschriebenen Art, so gehören die Strecken $p x^{(i)}$ wegen der Konvexität von M ganz zu M. Da nun die Längen dieser Strecken über alle Grenzen wachsen, ist also jeder Punkt eines Halbstrahls, gegen den sich die Halbstrahlen $\vec{p x}^{(i)}$ häufen, Häufungspunkt von Punkten aus M und gehört somit wegen der Abgeschlossenheit von M zu M.

Ferner läßt sich die Inklusion (11) für abgeschlossene konvexe M zu

$$\overline{B(M)} = B\big(A(M)\big) \tag{12}$$

verschärfen. Um dies einzusehen, bemerke man zunächst, daß es genügt, den Fall zu betrachten, wo $A(M)$ keinen vollständigen linearen Unterraum positiver Dimension enthält. Sei nämlich L_o der maximale in $A_o(M)$ enthaltene, durch den Nullpunkt o gehende lineare Unterraum, l seine Dimension und L_o^* der durch o gehende, zu L_o vollständig orthogonale $(n-l)$-dimensionale Unterraum. Der Polarkegel $B\big(A_o(M)\big)$ von $A_o(M)$ ist dann in L_o^* enthalten und hat als Teilmenge dieses Unterraums innere Punkte. Für jeden Punkt p

von M gilt wegen der genannten Eigenschaften der Kegel $A_p(M)$, daß der durch p gehende, zu L_o parallele l-dimensionale Unterraum ganz in M enthalten ist; mit anderen Worten, M ist die Vereinigung von zu L_o parallelen l-dimensionalen Unterräumen. Dies hat zur Folge, daß auch $B(M)$ in L_o^* enthalten ist und daß die Kegel $B(M)$ und $B\big(A_o(M)\big)$ ungeändert bleiben, wenn man statt E^n den Unterraum L_o^* betrachtet und M und somit $A_o(M)$ durch ihre Durchschnitte mit L_o^* ersetzt. Dies zeigt, daß wir ohne Beschränkung der Allgemeinheit zu den obigen Voraussetzungen hinzufügen können, daß $A(M)$ keinen linearen Unterraum positiver Dimension vollständig enthält. Es gibt dann Stützhyperebenen des Kegels $A_o(M)$, die mit ihm nur den Punkt o gemeinsam haben. Die positiven Normalen solcher Stützhyperebenen sind genau die inneren Halbstrahlen des Polarkegels $B\big(A_o(M)\big)$. Zum Beweis von (12) genügt es nun wegen (11) zu zeigen, daß ein nicht zu $B(M)$ gehöriger Halbstrahl nicht innerer Strahl von $B\big(A_o(M)\big)$ sein kann. Dies ist aber evident. Ist nämlich u ein nicht zu $B(M)$ gehöriger Vektor, so enthält der Halbraum $\sum u_i x_i > 0$ eine ins Unendliche strebende Folge von Punkten $x^{(i)}$ aus M. Die Häufungsstrahlen der Folge von Halbstrahlen $o\vec{x}^{(i)}$ sind folglich im Halbraum $\sum u_i x_i \geqq 0$ enthalten, und $A_o(M)$ kann somit (vom Scheitel o abgesehen) nicht vollständig dem offenen Halbraum $\sum u_i x_i < 0$ angehören.

Wir wenden uns nun zur Anwendung dieser Resultate auf die Niveaumannigfaltigkeiten konvexer Funktionen. Sei also $z = f(x)$ eine im offenen konvexen Bereich C definierte konvexe Funktion. Mit $[C, f]$ werde die abgeschlossene und konvexe Menge derjenigen Punkte (x, z) des E^{n+1} bezeichnet, für welche $x \in C$ und $z \geqq f(x)$. Es sei M_{z_0} der Durchschnitt von $[C, f]$ mit der Hyperebene $z = z_0$. Die orthogonale Projektion von M_{z_0} auf die Hyperebene $z = 0$ ist die Niveaumenge N_{z_0} von f, d.h. die Menge der Punkte x mit $f(x) \leqq z_0$. Der Asymptotenkegel $A_{p_0}[C, f]$, wo $p_0 = (x^{(0)}, z_0)$ ein Punkt von $[C, f]$ ist, wird von der Hyperebene $z = z_0$ im Asymptotenkegel $A_{p_0}(M_{z_0})$ von M_{z_0} geschnitten. Ist $p_1 = (x^{(1)}, z_1)$ ein zweiter Punkt von $[C, f]$, so geht $A_{p_1}[C, f]$ aus $A_{p_0}[C, f]$ durch eine Translation hervor. Dieselbe Translation führt aber die Hyperebene $z = z_0$ in $z = z_1$ über, und folglich ist $A(M_{z_0}) = A(M_{z_1})$. Projiziert man auf $z = 0$, so erhält man also die Behauptung:

Die Niveaumengen einer konvexen Funktion sind entweder alle beschränkt oder alle unbeschränkt mit demselben Asymptotenkegel.

Aus (12) kann man nun weiter schließen, daß die Kegel $B(N_z)$ der Beschränktheitsrichtungen für alle Niveaumengen N_z dieselbe abgeschlossene Hülle haben. Es läßt sich aber noch etwas mehr aussagen:

Hat eine konvexe Funktion $f(x)$ die untere Grenze a $(\geqq -\infty)$, so sind alle Niveaumengen N_z für $z > a$ in denselben Richtungen beschränkt. Ist $B = B(N_z)$ ihr gemeinsamer Kegel der Beschränktheitsrichtungen, so gilt, falls $f(x)$ die untere Grenze a erreicht, $B \subseteqq B(N_a) \subseteqq \bar{B}$.

Die zweite Behauptung folgt unmittelbar aus $N_a \subseteqq N_z$ für $z > a$ und dem soeben über die Kegel $B(N_z)$ Gesagten. Die Richtigkeit der ersten Behauptung kann man so einsehen. Es sei N_{z_0}, $z_0 > a$, in der Richtung u beschränkt.

Ist $\sum u_i x_i = c$ eine Schranke, so ist der $(n-1)$-dimensionale Unterraum L^{n-1} des E^{n+1} mit den Gleichungen $\sum u_i x_i = c$, $z = z_0$ eine Schranke der Menge M_{z_0} als Teilmenge der Hyperebene $z = z_0$. Dieser Unterraum L^{n-1} enthält also keine inneren Punkte der Menge $[C, f]$. Daher läßt sich durch ihn eine Schranke von $[C, f]$ legen. Durch Projektion auf eine zu L^{n-1} orthogonale (zweidimensionale) Ebene des E^{n+1} geht nämlich $[C, f]$ in einen (unbeschränkten) konvexen Bereich und L^{n-1} in einen äußeren oder Randpunkt dieses Bereichs über. Durch diesen Punkt läßt sich folglich eine Gerade legen, die Schranke des Bereichs ist und die daher zusammen mit L^{n-1} eine Schranke S von $[C, f]$ aufspannt. Nun muß S von der Hyperebene $z = z_0$ verschieden sein, da $[C, f]$ sowohl Punkte mit $z > z_0$ als auch Punkte mit $z < z_0$ enthält. Daher wird S von den Hyperebenen $z = \text{konst.}$ in zu L^{n-1} parallelen Schranken der Mengen M_z geschnitten. Projektion auf $z = 0$ ergibt also, daß jede Beschränktheitsrichtung von N_{z_0} auch Beschränktheitsrichtung aller übrigen Mengen N_z ist. Damit ist die Behauptung bewiesen.

Daß im Falle der Existenz eines Minimums von $f(x)$ der Kegel $B(N_a)$ der Menge N_a der Punkte, in denen das Minimum angenommen wird, umfassender sein kann als der Kegel B der Beschränktheitsrichtungen der übrigen Niveaumengen, zeigt die konvexe Funktion $f(x_1, x_2) = \sqrt{x_1^2 + x_2^2} - x_1$. Hier ist $a = 0$, und N_0 ist die positive x_1-Achse, also beschränkt in den Richtungen (u_1, u_2) mit $u_1 \leq 0$, während die anderen Mengen N_z von Parabeln mit der x_1-Achse als Achse begrenzt werden, also nur in Richtungen mit $u_1 < 0$ beschränkt sind.

Beispiele von Funktionen $\varphi(x_1, x_2)$, die die Bedingungen I bis III erfüllen, deren Niveaukurven aber nicht die soeben bewiesenen Eigenschaften der Niveaumannigfaltigkeiten konvexer Funktionen besitzen, sind ebenfalls leicht anzugeben. Die Niveaukurven der Funktion $\varphi(x_1, x_2) = (1 + x_2^2)/x_1^2$, $x_1 > 0$, sind die in $x_1 > 0$ gelegenen Zweige der konzentrischen Hyperbeln $t x_1^2 - x_2^2 = 1$, $t > 0$, mit von t abhängigem Asymptotenwinkel. Bei der Funktion $\varphi(x_1, x_2) = [(x_1 - 1)^2 + x_2^2]/x_1^2$, $x_1 > 0$, sind es die Kegelschnitte mit dem Brennpunkt $(1, 0)$ und der Leitlinie $x_1 = 0$. Die Schar enthält also hier sowohl beschränkte als auch unbeschränkte Kurven. In beiden Fällen ist die Bedingung IV nicht erfüllt.

Literatur

[1] COURANT, R., u. D. HILBERT: Methoden der mathematischen Physik, 2. Aufl., Bd. I. Berlin 1931. — [2] FENCHEL, W.: Convex cones, sets, and functions. (From notes by D. W. BLACKETT of lectures at Princeton University.) Princeton, N. J. 1953. — [3] FINETTI, B. DE: Sulle stratificazioni convesse. Ann. di Matematica (4) **30**, 173—183 (1949). — [4] KOWALEWSKI, G.: Einführung in die Determinantentheorie. Leipzig 1909. — [5] STEINITZ, E.: Bedingt konvergente Reihen und konvexe Systeme, II. J. reine u. angew. Math. **144**, 1—40 (1914). — [6] STOKER, J. J.: Unbounded convex point sets. Amer. J. Math. **62**, 165—179 (1940).

Kopenhagen, Dänische Technische Hochschule

(Eingegangen am 23. September 1955)

Kochendörffer, R.
Math. Zeitschr. Bd. 63, S. 507—513 (1956)

Über den Multiplikator einer Gruppe

Zum Gedenken an den 80. Geburtstag meines verehrten Lehrers I. Schur

Von

Rudolf Kochendörffer

In der von I. Schur[1]) begründeten Theorie der Darstellungen endlicher Gruppen durch gebrochene lineare Substitutionen wird jeder endlichen Gruppe $\mathfrak{H}$ ihr *Multiplikator* $\mathfrak{M} = \mathfrak{M}(\mathfrak{H})$ zugeordnet. Dieser spielt eine entscheidende Rolle bei der Aufgabe, eine Übersicht über alle Darstellungen von $\mathfrak{H}$ zu gewinnen. Der Multiplikator ist eine abelsche Gruppe, in deren Ordnung höchstens diejenigen Primzahlen aufgehen, die die Ordnung von $\mathfrak{H}$ teilen. Die vollständige Berechnung des Multiplikators einer gegebenen Gruppe ist in den meisten Fällen eine schwierige Aufgabe, so daß bereits Teilaussagen über seine Struktur von Interesse sein können. Eine solche Teilaussage ist die folgende:

Die p-Sylowgruppe des Multiplikators $\mathfrak{M}$ von $\mathfrak{H}$ ist isomorph zu einer Untergruppe des Multiplikators einer p-Sylowgruppe von $\mathfrak{H}$[2]).

Welche Untergruppe ihres Multiplikators eine Sylowgruppe in diesem Sinne zum Multiplikator der ganzen Gruppe beisteuert, hängt von ihrer Einbettung in die ganze Gruppe ab. Eine Bedingung dafür, daß eine abelsche Sylowgruppe überhaupt keinen Beitrag zum Multiplikator liefert, findet sich in S II[3]).

Im folgenden soll der oben formulierte Satz aufs neue bewiesen werden, und zwar mit einem Verfahren, welches eine Aussage darüber gestattet, wann eine Sylowgruppe ihren vollen Multiplikator zum Multiplikator der ganzen Gruppe beisteuert. Es ergibt sich folgender Satz:

Es sei $\mathfrak{S}$ eine p-Sylowgruppe von $\mathfrak{H}$ und

$$\mathfrak{H} = \sum_{V \in \mathfrak{V}} \mathfrak{S} V,$$

wobei V ein Rechtsvertretersystem $\mathfrak{V}$ von $\mathfrak{H}$ nach $\mathfrak{S}$ durchläuft. Besitzt der Komplex $\mathfrak{V}$ die Eigenschaft

$$P^{-1} \mathfrak{V} P = \mathfrak{V} \qquad \text{für jedes } P \text{ aus } \mathfrak{S},$$

so ist die p-Sylowgruppe des Multiplikators von $\mathfrak{H}$ isomorph zum Multiplikator von $\mathfrak{S}$.

[1]) I. Schur, Über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **127**, 20—50 (1904). — Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **132**, 85—137 (1907). Im folgenden zitiert mit SI und SII.

[2]) Vgl. SI, Satz IX.

[3]) SII, Seite 114.

Die Voraussetzung dieses Satzes ist insbesondere erfüllt, wenn $\mathfrak{S}$ direkter Faktor ist. In diesem Spezialfall folgt unsere Aussage aus einem Satz von SCHUR[4]).

Als Koeffizientenbereich für alle Substitutionen werde ein algebraisch abgeschlossener Körper zugrunde gelegt, dessen Charakteristik kein Teiler der Ordnung von $\mathfrak{H}$ ist. Ist

$$(A) = (a_{ik}) \qquad (i, k = 1, \ldots, n)$$

die Koeffizientenmatrix der dem Element A aus $\mathfrak{H}$ zugeordneten gebrochenen linearen Substitution

$$L_A: \qquad x_\nu = \frac{a_{\nu 1} y_1 + \cdots + a_{\nu, n-1} y_{n-1} + a_{\nu n}}{a_{n1} y_1 + \cdots + a_{n, n-1} y_{n-1} + a_{nn}} \qquad (\nu = 1, \ldots, n-1),$$

so gilt für je zwei Elemente A, B aus $\mathfrak{H}$

$$(1) \qquad (A)\,(B) = r_{A,B}(AB)$$

mit einem gewissen Faktorensystem $\{r_{A,B}\}$ aus nicht verschwindenden Zahlen. Infolge des assoziativen Gesetzes in $\mathfrak{H}$ genügen die Zahlen $r_{A,B}$ den Relationen

$$(2) \qquad r_{A,BC}\, r_{B,C} = r_{AB,C}\, r_{A,B}.$$

Umgekehrt gehören zu jedem Faktorensystem $\{r_{A,B}\}$, welches die Bedingungen (2) erfüllt, Matrizen $(A), (B), \ldots$, die den Gln. (1) genügen und somit eine Darstellung von $\mathfrak{H}$ durch gebrochene lineare Substitutionen bilden.

Ersetzt man die Matrizen (A) durch proportionale Matrizen $c_A \cdot (A)$ mit beliebigen Zahlfaktoren $c_A \neq 0$, so erhält man die gleichen gebrochenen linearen Substitutionen, während das Faktorensystem $\{r_{A,B}\}$ in das System

$$(3) \qquad \{r'_{A,B}\} = \left\{\frac{c_A c_B}{c_{AB}} \cdot r_{A,B}\right\}$$

übergeht. Zwei Faktorensysteme $\{r_{A,B}\}$ und $\{r'_{A,B}\}$, die durch (3) miteinander zusammenhängen, heißen assoziiert, in Zeichen:

$$\{r_{A,B}\} \sim \{r'_{A,B}\}.$$

Ein Faktorensystem, welches zum Einssystem $\{1\}$ assoziiert ist, wird als ein zerfallendes Faktorensystem bezeichnet.

Hat man zwei Faktorensysteme $\{r_{A,B}\}$ und $\{s_{A,B}\}$, die (2) erfüllen, so genügen offenbar auch das System $\{r^{-1}_{A,B}\}$ sowie das Produktsystem

$$\{r_{A,B}\}\{s_{A,B}\} = \{r_{A,B} \cdot s_{A,B}\}$$

den Bedingungen (2) und gehören infolgedessen ebenfalls zu Darstellungen von $\mathfrak{H}$ durch gebrochene lineare Substitutionen. Da die Assoziiertheit eine mit dieser Multiplikation der Faktorensysteme verträgliche Äquivalenzrelation darstellt, bilden die Klassen assoziierter Faktorensysteme eine abelsche Gruppe

[4]) SII, Satz VII.

$\mathfrak{M}$, die als der *Multiplikator* von $\mathfrak{H}$ bezeichnet wird. Es zeigt sich, daß $\mathfrak{M}$ eine endliche Ordnung besitzt, in der nur solche Primzahlen aufgehen, die die Ordnung von $\mathfrak{H}$ teilen.

Die Grundlage für das Folgende bildet eine Schlußweise von W. GASCHÜTZ[5]), die sich bei der Untersuchung von Gruppenerweiterungen bewährt hat.

Für die Zerlegung von $\mathfrak{H}$ in Rechtsnebenklassen nach $\mathfrak{S}$ legen wir ein festes Vertretersystem $\mathfrak{V}$ zugrunde und haben dann

$$\mathfrak{H} = \sum_{V \in \mathfrak{V}} \mathfrak{S} V. \tag{4}$$

Sei $\{r'_{A,B}\}$ irgendein zur Gruppe $\mathfrak{H}$ gehöriges Faktorensystem. Das Teilsystem $\{r'_{P,Q}\}$ für die Elementepaare $P, Q \in \mathfrak{S}$ bildet dann ein Faktorensystem für die Sylowgruppe $\mathfrak{S}$ von $\mathfrak{H}$. Wir wollen nun zunächst das Faktorensystem $\{r'_{A,B}\}$ durch ein assoziiertes System

$$\{r_{A,B}\} = \left\{ \frac{c_{AB}}{c_A c_B} \cdot r'_{A,B} \right\}$$

ersetzen, und zwar durch folgende Wahl der Zahlen c_A:

$$c_P = 1 \quad \text{für} \quad P \in \mathfrak{S},$$

$$c_V \neq 0, \text{ beliebig für } V \in \mathfrak{V},$$

$$c_{PV} = c_P c_V r'^{-1}_{P,V}.$$

Dadurch wird erreicht, daß das Teilfaktorensystem $\{r'_{P,Q}\}$ mit $P, Q \in \mathfrak{S}$ ungeändert bleibt, während für $P \in \mathfrak{S}$ und $V \in \mathfrak{V}$

$$r_{P,V} = \frac{c_{PV}}{c_P c_V} \cdot r'_{P,V} = \frac{c_P c_V r'^{-1}_{P,V}}{c_P c_V} r'_{P,V} = 1 \tag{5}$$

wird.

Um die jetzt vorzunehmenden Umformungen übersichtlich zu gestalten, empfiehlt es sich, folgende Bezeichnungsweise aus [5]) zu übernehmen:

Ein beliebiges Element H aus $\mathfrak{H}$ besitzt gemäß (4) eine eindeutige Zerlegung

$$H = PV \quad \text{mit} \quad P \in \mathfrak{S}, \; V \in \mathfrak{V}.$$

Wir setzen

$$P = \underline{H}, \; V = \overline{H}, \quad \text{also} \quad H = \underline{H}\overline{H}.$$

Die Relationen (2) benutzen wir im Spezialfall $A = P \in \mathfrak{S}$, $B = \underline{H}$, $C = \overline{H}$ und erhalten

$$r_{P,\underline{H}\overline{H}} r_{\underline{H},\overline{H}} = r_{P\underline{H},\overline{H}} r_{P,\underline{H}},$$

also wegen (5)

$$r_{P,H} = r_{P,\underline{H}\overline{H}} = r_{P,\underline{H}} \qquad (P \in \mathfrak{S}, \; H \in \mathfrak{H}). \tag{6}$$

Ein zweites Mal benutzen wir (2) mit $A = \underline{H}$, $B = \overline{H}$, $C \in \mathfrak{H}$:

$$r_{\underline{H},\overline{H}C} r_{\overline{H},C} = r_{\underline{H}\overline{H},C} r_{\underline{H},\overline{H}} = r_{H,C} r_{\underline{H},\overline{H}}.$$

[5]) GASCHÜTZ, W.: Zur Erweiterungstheorie der endlichen Gruppen. J. reine u. angew. Math. **190**, 93—107 (1952).

Hieraus folgt unter Berücksichtigung von (5)

$$r_{\underline{H},\overline{HC}}\, r_{\overline{H},C} = r_{H,C}$$

und weiter, indem man (6) auf den ersten Faktor der linken Seite anwendet,

$$r_{\underline{H},\underline{\overline{HC}}}\, r_{\overline{H},C} = r_{H,C} \qquad (C, H \in \mathfrak{H}). \tag{7}$$

Die Relation (2) läßt sich in der Form

$$r_{B,C} = r^{-1}_{A,BC}\, r_{A,B}\, r_{AB,C}$$

schreiben. Den letzten Faktor rechts formen wir mittels (7) um und erhalten

$$r_{B,C} = r^{-1}_{A,BC}\, r_{A,B}\, r_{\overline{AB},C}\, r_{\underline{AB},\underline{\overline{ABC}}}.$$

Hierin lasse man A die $k = [\mathfrak{H}:\mathfrak{S}]$ Elemente des Vertretersystems $\mathfrak{V}$ durchlaufen und multipliziere alle diese Gleichungen miteinander. Das ergibt

$$r^k_{B,C} = \prod_{V\in\mathfrak{V}} r^{-1}_{V,BC} \cdot \prod_{V\in\mathfrak{V}} r_{V,B} \cdot \prod_{V\in\mathfrak{V}} r_{V,C} \cdot \prod_{V\in\mathfrak{V}} r_{\underline{VB},\underline{\overline{VB}C}},$$

wobei man für die dritten Faktoren auf der rechten Seite nur zu beachten hat, daß $\overline{VB}$ bei festem B zugleich mit V alle Elemente aus $\mathfrak{V}$ durchläuft. Setzt man zur Abkürzung

$$\prod_{V\in\mathfrak{V}} r_{V,H} = a_H,$$

so erhält man

$$r^k_{B,C} = \frac{a_B\, a_C}{a_{BC}} \cdot \prod_{V\in\mathfrak{V}} r_{\underline{VB},\underline{\overline{VB}C}}. \tag{8}$$

Das Zahlensystem

$$\{s_{B,C}\} = \{\prod_{V\in\mathfrak{V}} r_{\underline{VB},\underline{\overline{VB}C}}\} \tag{9}$$

hat sich mithin als ein zu $\{r^k_{B,C}\}$ assoziiertes Faktorensystem erwiesen.

Jedem zur vollen Gruppe $\mathfrak{H}$ gehörigen Faktorensystem

$$\{r_{A,B}\} \qquad (A, B \in \mathfrak{H})$$

läßt sich ein zur Untergruppe $\mathfrak{S}$ gehöriges Faktorensystem zuordnen, nämlich einfach das Teilsystem

$$\{r_{P,Q}\} \qquad (P, Q \in \mathfrak{S}).$$

Assoziierten $\mathfrak{H}$-Faktorensystemen entsprechen bei dieser Zuordnung selbstverständlich assoziierte $\mathfrak{S}$-Faktorensysteme. Das Produkt zweier $\mathfrak{H}$-Faktorensysteme wird abgebildet auf das Produkt der zu $\mathfrak{S}$ gehörigen Teilsysteme. Man erhält mithin eine homomorphe Abbildung des Multiplikators $\mathfrak{M}$ von $\mathfrak{H}$ auf eine Untergruppe $\mathfrak{U}$ des Multiplikators $\mathfrak{N}$ von $\mathfrak{S}$. Wir bestimmen den Kern dieses Homomorphismus. Da die Ordnung von $\mathfrak{N}$, wie oben bemerkt wurde, eine Potenz von p ist, gehören sicher alle diejenigen Elemente von $\mathfrak{M}$ zum Kern, deren Ordnungen zu p teilerfremd sind. Folglich brauchen nur noch die Elemente der p-Sylowgruppe $\mathfrak{P}$ von $\mathfrak{M}$ weiter untersucht zu werden.

Es wird sich zeigen, daß kein von Eins verschiedenes Element aus $\mathfrak{P}$ zum Kern des betrachteten Homomorphismus gehört oder, mit anderen Worten, daß für die Faktorensysteme aus den $\mathfrak{P}$ angehörenden Klassen aus dem Zerfallen des zu $\mathfrak{S}$ gehörenden Teilsystems bereits das Zerfallen des gesamten Systems folgt.

Um den Nachweis hierfür zu erbringen, kann (8) dienen, weil in dem Produkt auf der rechten Seite nur Zahlen aus dem zu $\mathfrak{S}$ gehörigen Teilsystem auftreten. Sei also $\{r_{A,B}\}$ ein $\mathfrak{H}$-Faktorensystem aus einer der zu $\mathfrak{P}$ gehörenden Klassen, für welches das der Untergruppe $\mathfrak{S}$ entsprechende Teilsystem zerfällt, d.h.

$$r_{P,Q} = \frac{d_P d_Q}{d_{PQ}} \qquad (P, Q \in \mathfrak{S}).$$

Für das Zahlensystem (9) erhält man dann

$$\prod_{V \in \mathfrak{V}} r_{\underline{VB}, \overline{VB}C} = \frac{\prod\limits_{V \in \mathfrak{V}} d_{\underline{VB}} \cdot \prod\limits_{V \in \mathfrak{V}} d_{\overline{VB}C}}{\prod\limits_{V \in \mathfrak{V}} d_{\underline{VB}\,\overline{VB}C}}.$$

Mit der Abkürzung

$$\prod_{V \in \mathfrak{V}} d_{\underline{VB}} = e_B$$

wird

$$\prod_{V \in \mathfrak{V}} d_{\overline{VB}C} = \prod_{V \in \mathfrak{V}} d_{\underline{VC}} = e_C$$

und

$$\prod_{V \in \mathfrak{V}} d_{\underline{VB}\,\overline{VB}C} = \prod_{V \in \mathfrak{V}} d_{\underline{VBC}} = e_{BC}.$$

Bei der letzten Umformung ist von der Gleichung

(10) $$\underline{VB}\,\overline{\underline{VB}C} = \underline{VBC}$$

Gebrauch gemacht worden, die man leicht folgendermaßen bestätigt: Es werde gesetzt:

$$B = S_1 V_1, \quad C = S_2 V_2 \qquad (S_1, S_2 \in \mathfrak{S};\ V_1, V_2 \in \mathfrak{V}).$$

Dann wird

$$VB = V S_1 V_1 = S_1' V' V_1 = S_1' S W$$

mit

$$V S_1 = S_1' V' \quad \text{und} \quad V' V_1 = S W \qquad (S_1',\ S \in \mathfrak{S};\ V', W \in \mathfrak{V}).$$

Folglich ist

$$\underline{VB} = S_1' S, \quad \overline{VB} = W.$$

Ferner erhält man

$$\overline{\underline{VB}C} = \underline{W S_2 V_2} = \underline{S_2' W' V_2} = \underline{S_2' \overline{S}\,\overline{W}} = S_2' \overline{S},$$

wenn

$$W S_2 = S_2' W' \quad \text{und} \quad W' V_2 = \overline{S}\,\overline{W} \qquad (S_2', \overline{S} \in \mathfrak{S};\ W', \overline{W} \in \mathfrak{V})$$

gesetzt wird. Zusammen ist also

$$\underline{VB}\,\overline{\underline{VB}C} = S_1' S S_2' \overline{S}.$$

Andererseits ist

$$VBC = \underline{V S_1 V_1 S_2 V_2} = \underline{S_1' V' V_1 S_2 V_2} = \underline{S_1' S W S_2 V_2}$$
$$= \underline{S_1' S S_2' W' V_2} = \underline{S_1' S S_2' \overline{S}\,\overline{W}} = S_1' S S_2' \overline{S},$$

also in der Tat

$$\underline{VB}\,\underline{\overline{VB}C} = \underline{VBC}.$$

Im ganzen erhalten wir daher für die Zahlen des Systems (9)

$$\prod_{V \in \mathfrak{V}} r_{\underline{VB}, \underline{\overline{VB}C}} = \frac{e_B e_C}{e_{BC}}.$$

Daher besagt (8), daß das System $\{r_{A,B}\}^k$ zerfällt. Da der Exponent $k = [\mathfrak{H} : \mathfrak{S}]$ zur Ordnung der das Faktorensystem $\{r_{A,B}\}$ enthaltenden Klasse aus $\mathfrak{P}$ teilerfremd ist, zerfällt $\{r_{A,B}\}$ selbst.

Mithin hat sich ergeben, daß $\mathfrak{P}$ isomorph einer Untergruppe $\mathfrak{U}$ des Multiplikators von $\mathfrak{S}$ ist. Damit ist der erste der oben genannten Sätze bewiesen.

Die Untergruppe $\mathfrak{U}$ des Multiplikators von $\mathfrak{S}$ kann folgendermaßen gekennzeichnet werden: $\mathfrak{U}$ besteht genau aus allen denjenigen Klassen von $\mathfrak{S}$-Faktorensystemen, die als Teilsysteme von $\mathfrak{H}$-Faktorensystemen aufgefaßt werden können oder, wie wir auch sagen wollen, die sich zu $\mathfrak{H}$-Faktorensystemen fortsetzen lassen. Formel (8) lehrt, daß die Fortsetzung eines $\mathfrak{S}$-Faktorensystems zu einem $\mathfrak{H}$-Faktorensystem — falls überhaupt — bis auf assoziierte Systeme nur auf eine einzige Weise möglich ist. Denn sind $\{q_{A,B}\}$ und $\{q'_{A,B}\}$ zwei Fortsetzungen ein und desselben $\mathfrak{S}$-Faktorensystems, so zerfällt von dem System $\{q'_{A,B} \cdot q_{A,B}^{-1}\}$ der zu $\mathfrak{S}$ gehörige Teil, woraus man wie oben schließt, daß das gesamte System zerfällt. Wie die Fortsetzung eines $\mathfrak{S}$-Faktorensystems aussehen muß, falls sie überhaupt existiert, ist ebenfalls aus (8) zu ersehen. Man bestimme eine Lösung l der Kongruenz

$$kl \equiv 1 \pmod{[\mathfrak{P} : 1]},$$

was stets möglich ist, da $k = [\mathfrak{H} : \mathfrak{S}]$ zu der p-Potenz $[\mathfrak{P} : 1]$ teilerfremd ist. Dann folgt aus (8), daß das Faktorensystem $\{r_{B,C}\}$ zu dem System

$$\{t_{B,C}\} = \{s_{B,C}\}^l = \left\{\prod_{V \in \mathfrak{V}} r_{\underline{VB}, \underline{\overline{VB}C}}\right\}^l$$

assoziiert ist. Falls also das $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ überhaupt eine Fortsetzung besitzt, so muß diese zu $\{t_{B,C}\}$ assoziiert sein.

Das Zahlensystem $\{t_{B,C}\}$ läßt sich nun für jedes beliebige $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ bilden. Man bestätigt ohne Mühe, daß $\{t_{B,C}\}$ stets ein $\mathfrak{H}$-Faktorensystem darstellt, indem man die Richtigkeit der Gleichungen

$$t_{A,BC}\, t_{B,C} = t_{AB,C}\, t_{A,B}$$

nachprüft. Dabei hat man neben (10) von der Beziehung

$$\overline{\overline{VA}\,B} = \overline{VAB}$$

Gebrauch zu machen, die man ähnlich wie (10) leicht beweist. Hat man für ein beliebiges $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ das System $\{t_{B,C}\}$ gebildet, so bleibt zu untersuchen, ob dadurch tatsächlich eine Fortsetzung von $\{r_{P,Q}\}$ gewonnen ist, d.h. ob die beiden $\mathfrak{S}$-Faktorensysteme $\{r_{P,Q}\}$ und $\{t_{P,Q}\}$ assoziiert sind. Falls jedes $\mathfrak{S}$-Faktorensystem $\{r_{P,Q}\}$ zu dem entsprechenden System $\{t_{P,Q}\}$ assoziiert ist, kann jedes $\mathfrak{S}$-Faktorensystem zu einem $\mathfrak{H}$-Faktorensystem fortgesetzt werden. Die Beziehung

$$\{r_{P,Q}\} \sim \{t_{P,Q}\}$$

ist gleichbedeutend mit

$$\{r_{P,Q}\}^k \sim \{t_{P,Q}\}^k$$

und wegen

$$\{t_{P,Q}\}^k \sim \{s_{P,Q}\} = \left\{ \prod_{V \in \mathfrak{V}} r_{\underline{VP}, \underline{\overline{V}\overline{P}Q}} \right\}$$

auch gleichbedeutend mit dem Zerfallen des Systems

$$\left\{ \prod_{V \in \mathfrak{V}} \frac{r_{\underline{VP}, \underline{\overline{V}\overline{P}Q}}}{r_{P,Q}} \right\}.$$

Dieses System zerfällt sicher, wenn jeder einzelne Faktor des Produktes gleich 1 ist, d.h.

$$r_{\underline{VP}, \underline{\overline{V}\overline{P}Q}} = r_{P,Q} \qquad (P, Q \in \mathfrak{S};\ V \in \mathfrak{V}).$$

Das tritt jedenfalls ein, wenn für jedes Paar $P \in \mathfrak{S}$, $V \in \mathfrak{V}$

$$\underline{VP} = P$$

wird oder

$$VP = PP^{-1}VP = PV'$$

mit $V' = P^{-1}VP \in \mathfrak{V}$. Das bedeutet, daß sich ein Vertretersystem $\mathfrak{V}$ wählen läßt, welches die Eigenschaft

$$P^{-1}\mathfrak{V}P = \mathfrak{V} \qquad \text{für jedes } P \in \mathfrak{S}$$

besitzt. Damit ist der zweite der oben angegebenen Sätze bewiesen.

Rostock, Mathematisches Institut der Universität

(Eingegangen am 9. August 1955)

Geiringer, H.
Math. Zeitschr. Bd. 63, S. 514—524 (1956)

Grenzlinien der Hodographentransformation

Issai Schur zum Gedächtnis

Von

Hilda Geiringer

In der Theorie der ebenen Hodographentransformation der Aerodynamik spielen gewisse Singularitäten dieser Transformation, die dem Verschwinden Jacobischer Funktionaldeterminanten entsprechen, eine wichtige Rolle. Es ist verwunderlich, daß die Diskussion solcher Singularitäten nicht von einem allgemeinen Standpunkt aus in Lehrbüchern der Differential- und Integralrechnung Platz findet. Diese beschränken sich im allgemeinen auf die Bemerkung, daß die Funktionaldeterminante der Abbildung als von Null verschieden vorausgesetzt wird; es wird aber kaum darauf eingegangen, was im Falle ihres Verschwindens geschieht. Vorbereitet durch Arbeiten von Ringleb [1] und Tollmien [2] finden sich Studien dieser Richtung, mit speziellem Bezug auf das aerodynamische Problem, unter anderem in einer inhaltsreichen Arbeit von Craggs [3], in dem bekannten Lehrbuch von Courant und Friedrichs [4] und in dem von Howarth [5] herausgegebenen Sammelwerk.

Die folgende Darstellung unterscheidet sich von den genannten teils in der Form, teils im Inhalt. Unsere sehr einfache Methode scheint den Methoden der unter [3], [4], [5] genannten Darstellungen überlegen. Auch inhaltlich gehen wir über die in diesen Arbeiten gegebenen Resultate (besonders über die in [4] gegebenen) in mancher Hinsicht hinaus. Die hier benutzte Methode wurde in einer mir erst nachträglich bekannt gewordenen Arbeit von R. E. Meyer [6] angewendet, die sich aber nicht mit der Darstellung der grundlegenden Eigenschaften der Singularitäten, sondern mit verfeinerten Fragestellungen vom Standpunkt der Riemannschen Geometrie beschäftigt. Hingegen ist die fragliche Methode schon 1941 von J. Mandel [7] angewendet worden, allerdings nicht auf die Hodographentransformation der Aerodynamik, sondern auf das Spannungsproblem der ebenen Plastizitätstheorie. Dieses hat mit jener mathematisch entscheidende Ähnlichkeiten [8], [9]. Die vorliegende Darstellung ist von der Mandels methodisch beeinflußt.

Um Raum zu sparen, geben wir unsere Resultate gleich in aerodynamischer Form, ohne erst auf die allgemeine Abbildungsaufgabe einzugehen.

1. Aerodynamische Grundgleichungen. Wir betrachten das ebene stationäre Problem eines kompressiblen, wirbelfreien, reibungsfreien Gases. Dieses wird als „elastisch" angenommen, d. h. es existiert eine universale, eineindeutige Beziehung zwischen Druck p und Dichte ϱ. Seien x, y die cartesischen Koordinaten der Strömungsebene, u, v die x- und y-Komponenten des Geschwindigkeitsvektors $\bar{q}$, dessen Polarkoordinaten q und ϑ sind. Die

u, v-Ebene heißt Hodographenebene oder Hodograph. Bezeichnet $\frac{\partial}{\partial s}$ und $\frac{\partial}{\partial n}$ Differentiation in Richtung von $\bar{q}$ und normal dazu, so ist

$$(1) \qquad \frac{\partial}{\partial s} = \cos\vartheta \frac{\partial}{\partial x} + \sin\vartheta \frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial n} = -\sin\vartheta \frac{\partial}{\partial x} + \cos\vartheta \frac{\partial}{\partial y},$$

und es gelten die Gleichungen

$$(2) \qquad \frac{\partial q}{\partial n} = q \frac{\partial \vartheta}{\partial s}, \qquad (M^2 - 1) \frac{\partial q}{\partial s} = q \frac{\partial \vartheta}{\partial n},$$

die Wirbelfreiheit und Kontinuität ausdrücken. Dabei ist

$$(3) \qquad M^2 = q^2 \frac{d\varrho}{dp}$$

und wir setzen $\frac{dp}{d\varrho} = a^2$; ferner ist q^2 mit p und ϱ durch die Bernoullische Gleichung

$$(4) \qquad \frac{q^2}{2} + \int^{p} \frac{dp}{\varrho} = \text{constant}$$

verbunden. Mittels (3) und (4) kann M^2 als Funktion von q^2 ausgedrückt werden. Wird die p, ϱ-Beziehung in der adiabatischen Form $p\varrho^{-\gamma} = \text{constant}$ angenommen, so folgt

$$(3') \qquad M^2 = \frac{q^2}{a^2} = \frac{q^2}{a_0^2 - \frac{\gamma - 1}{2} q^2},$$

wo a_0 den Wert der Schallgeschwindigkeit a für $q = 0$ („Stagnationswert" von a) bezeichnet. M^2 wächst monoton mit q^2 von $M = 0$ für $q = 0$, zu $M = \infty$ für die Maximalgeschwindigkeit q_m, für welche p, ϱ und a verschwinden; für $M = 1$ wird die Schallgeschwindigkeit oder „kritische" Geschwindigkeit $q = a$ angenommen. Ist M^2 durch q^2 ausgedrückt, so hat man in (2) zwei Gleichungen für q und ϑ, und p und ϱ folgen sodann mittels (4) und der p, ϱ-Beziehung.

Für $M > 1$ (Überschallgeschwindigkeit) setzt man

$$(5) \qquad M = \frac{1}{\sin\alpha}, \qquad M^2 - 1 = \operatorname{ctg}^2\alpha, \qquad 0 \leq \alpha \leq 90^\circ$$

und es ergibt sich, daß

$$(6) \qquad \varphi^+ = \vartheta + \alpha, \qquad \varphi^- = \vartheta - \alpha$$

die *charakteristischen Richtungen* von (2) sind. Die überall den Winkel φ^- mit der x-Achse einschließenden Kurven werden als C^--Charakteristiken oder als ξ-Kurven, die andern als C^+-Charakteristiken oder η-Kurven bezeichnet. Setzt man

$$(7) \qquad \frac{dQ}{dq} = \frac{1}{q \operatorname{tang}\alpha}, \qquad Q = \int^{q} \frac{dq}{q \operatorname{tang}\alpha},$$

so hat $Q + \vartheta$ entlang einer C^-, und $Q - \vartheta$ entlang einer C^+ konstanten Wert *(Kompatibilitätsbedingungen)*. Wir können daher

$$(8) \qquad Q + \vartheta = 2\eta, \qquad Q - \vartheta = 2\xi,$$

$$(9) \qquad Q = \eta + \xi, \qquad \vartheta = \eta - \xi$$

setzen.

Die Gln. (2) sind *nichtlinear*, da die Koeffizienten von der abhängigen Variabeln q abhängen. Da sie aber *homogen* sind und die Koeffizienten *nur* die abhängigen Variabeln enthalten (solche Gleichungen nennen wir reduzibel), so kann (2) *durch einfache Vertauschung von* x, y *mit* q, ϑ in ein System von *linearen* Gleichungen mit den unabhängigen Variabeln q, ϑ transformiert werden.

Führt man das *Potential* φ und die *Stromfunktion* ψ ein, wo $u = \frac{\partial \varphi}{\partial x}$, $v = \frac{\partial \varphi}{\partial y}$ und

$$\frac{\partial \varphi}{\partial x} = \frac{1}{\varrho} \frac{\partial \psi}{\partial y}, \qquad \frac{\partial \varphi}{\partial y} = -\frac{1}{\varrho} \frac{\partial \psi}{\partial x}, \tag{9}$$

so stellt (9) wieder ein System nichtlinearer reduzibler Gleichungen dar, da ϱ von $q = \left[\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial y}\right)^2\right]^{\frac{1}{2}}$ abhängt. Betrachtet man nun wieder φ und ψ als Funktionen nicht von x, y, sondern von q, ϑ (dies geht auf Chaplygin zurück und entspricht der oben erwähnten Vertauschung), so erhält man die fundamentalen *linearen* Gleichungen

$$\frac{\partial \varphi}{\partial \vartheta} = \frac{q}{\varrho} \frac{\partial \psi}{\partial q}, \quad (M^2 - 1) \frac{\partial \psi}{\partial \vartheta} = \varrho\, q \frac{\partial \varphi}{\partial q} \tag{10}$$

für φ, ψ als abhängige und q, ϑ als unabhängige Variable.

2. Grenzkurven und Verzweigungskurven. Hilfsformeln. Die Einführung von u, v als neuen unabhängigen und x, y als neuen abhängigen Variablen (Übergang zur Hodographenebene) ist nur möglich, wenn

$$j = \frac{\partial (u, v)}{\partial (x, y)} \neq 0 \tag{11}$$

ist. Man erhält durch diese Hodographentransformation

$$\frac{\partial x}{\partial u} = \frac{1}{j} \frac{\partial v}{\partial y}, \quad \frac{\partial x}{\partial v} = -\frac{1}{j} \frac{\partial u}{\partial y}, \quad \frac{\partial y}{\partial u} = -\frac{1}{j} \frac{\partial v}{\partial x}, \quad \frac{\partial y}{\partial v} = \frac{1}{j} \frac{\partial u}{\partial x}. \tag{12}$$

Damit man anderseits von einer „Hodographenlösung" $\bar{r} = \bar{r}(\bar{q})$ zu einer „physikalischen" Lösung $\bar{q} = \bar{q}(\bar{r})$ zurückkehren kann, muß mit $I j = 1$

$$I = \frac{\partial (x, y)}{\partial (u, v)} \neq 0 \tag{13}$$

sein. Linien, längs denen (13) nicht gilt, werden als *Grenzlinien* der „physikalischen" x, y-Ebene bezeichnet, solche, längs denen (11) versagt, als Verzweigungslinien dieser Ebene[1]).

Wir beginnen mit einigen *Hilfsformeln*. Wir bezeichnen mit ds_1, ds_2 das Linienelement einer ξ- oder η-Linie und setzen

$$h_1 = \frac{ds_1}{d\xi}, \quad h_2 = \frac{ds_2}{d\eta}. \tag{14}$$

[1]) Für eine sog. PRANDTL-MEYER-Strömung ist $j = 0$ in einem zweidimensionalen Gebiet.

Dann folgt für den Vektor $d\bar{r}$ mit den Komponenten dx, dy

$$\frac{\partial \bar{r}}{\partial \xi} = h_1 \bar{\xi}^0, \qquad \frac{\partial \bar{r}}{\partial \eta} = h_2 \bar{\eta}^0, \tag{15}$$

wenn $\bar{\xi}^0$ und $\bar{\eta}^0$ Einheitsvektoren in der $(\vartheta - \alpha)$- und $(\vartheta + \alpha)$-Richtung sind, und mit der Bezeichnung (6):

$$\begin{cases} \dfrac{\partial x}{\partial \xi} = h_1 \cos \varphi^-, & \dfrac{\partial x}{\partial \eta} = h_2 \cos \varphi^+ \\ \dfrac{\partial y}{\partial \xi} = h_1 \sin \varphi^-, & \dfrac{\partial y}{\partial \eta} = h_2 \sin \varphi^+ \end{cases} \tag{15'}$$

und

$$J = \frac{\partial (x, y)}{\partial (\xi, \eta)} = h_1 h_2 \sin 2\alpha. \tag{16}$$

Diese der Abbildung der physikalischen, d.h. der x, y-Ebene auf die charakteristische Ebene entsprechende Funktionaldeterminante werden wir im Überschallgebiet vorwiegend benutzen.

Schreiben wir von nun an kurz $\frac{\partial \varphi}{\partial x} = \varphi_x$, $\frac{\partial \psi}{\partial \xi} = \psi_\xi$ usw., so bestätigt man leicht, daß

$$\begin{cases} \varphi_\xi = h_1 \cdot q \cos \alpha, & \varphi_\eta = h_2 q \cos \alpha \\ \psi_\xi = -h_1 \varrho q \sin \alpha, & \psi_\eta = h_2 \varrho q \sin \alpha \end{cases} \tag{17}$$

ist, und für $D = \frac{\partial (\varphi, \psi)}{\partial (\varrho, \vartheta)}$ findet man (z.B. nach (10) und (7))

$$\begin{cases} D = \dfrac{\operatorname{ctg}^2 \alpha}{\varrho q} (\psi_\vartheta^2 - \psi_Q^2) = -\dfrac{\operatorname{ctg}^2 \alpha}{\varrho q} \psi_\xi \psi_\eta = \dfrac{\varrho}{q} \varphi_\xi \varphi_\eta \\ \qquad = h_1 h_2 \varrho q \cos^2 \alpha = \dfrac{\varrho q}{2} \operatorname{ctg} \alpha \cdot J. \end{cases} \tag{18}$$

Für die in üblicher Weise definierten *Krümmungen* $\varkappa_1, \varkappa_2, \varkappa$ und Krümmungsradien R_1, R_2, R der charakteristischen Kurven und der Stromlinien erhält man

$$\varkappa_1 = \frac{1}{R_1} = \frac{\partial (\vartheta - \alpha)}{\partial s_1} = \frac{\partial (\vartheta - \alpha)}{h_1 \partial \xi}, \qquad \varkappa_2 = \frac{1}{R_2} = -\frac{\partial (\vartheta + \alpha)}{\partial s_2} = -\frac{\partial (\vartheta + \alpha)}{h_2 \partial \eta}, \tag{19}$$

$$\varkappa = \frac{\partial \vartheta}{\partial s} = \frac{1}{2 \cos \alpha} \left(\frac{\partial \vartheta}{\partial s_1} + \frac{\partial \vartheta}{\partial s_2} \right) = \frac{1}{2 \cos \alpha} \left(\frac{1}{h_2} - \frac{1}{h_1} \right). \tag{19'}$$

Wir berechnen in gleicher Weise die *Beschleunigung* b und finden:

$$b = \frac{dq}{dt} = q \frac{\partial q}{\partial s} = \frac{q}{2 \cos \alpha} \left(\frac{\partial q}{\partial s_1} + \frac{\partial q}{\partial s_2} \right) = \frac{q^2 \sin \alpha}{2 \cos^2 \alpha} \left(\frac{1}{h_1} + \frac{1}{h_2} \right). \tag{20}$$

Wenn man mit $'$ Differentiation nach q bezeichnet und (8) benutzt, so folgt

$$h_1 = -R_1 \left(1 + \frac{\alpha'}{Q'} \right), \qquad h_2 = -R_2 \left(1 + \frac{\alpha'}{Q'} \right), \qquad h_1 / h_2 = R_1 / R_2. \tag{21}$$

Setzt man die zweiten gemischten Ableitungen von (15') einander gleich, und subtrahiert die so erhaltenen Gleichungen, so folgt, nach Vereinfachung, das

System *linearer* Gleichungen für h_1, h_2:

$$(22)\qquad \begin{cases} \frac{\partial h_1}{\partial \eta}\sin 2\alpha + \left(\frac{\alpha'}{Q'} - 1\right)(h_2 + h_1 \cos 2\alpha) = 0 \\ \frac{\partial h_2}{\partial \xi}\sin 2\alpha + \left(\frac{\alpha'}{Q'} - 1\right)(h_1 + h_2 \cos 2\alpha) = 0, \end{cases}$$

die für die folgende Diskussion benutzt werden.

3. Fall der Unterschallgeschwindigkeit. Wenn $M < 1$ ist in

$$(18')\qquad D = \frac{\partial(\varphi, \psi)}{\partial(q, v)} = -\frac{1}{\varrho q}\left[q^2 \psi_q^2 + (1 - M^2)\psi_\vartheta^2\right],$$

so verschwindet D nur, wenn beide Summanden in der Klammer verschwinden, d.h. wenn

$$q\psi_q = 0 \quad \text{und} \quad \psi_\vartheta = 0,$$

und nach (10)

$$\varphi_\vartheta = 0, \qquad q\varphi_q = 0.$$

Also ist dann entweder $q = 0$ (Stagnationspunkt) und außerdem $\psi_\vartheta = \varphi_\vartheta = 0$; oder es ist $q \neq 0$, dann *müssen alle vier Ableitungen* $\psi_q, \psi_\vartheta, \varphi_q, \varphi_\vartheta$ *und daher auch* $x_q, x_\vartheta, y_q, y_\vartheta$ *verschwinden.* (Man kann im letzteren Falle leicht sehen, daß der singuläre Punkt ein *Sattelpunkt von* $\varphi(q, \vartheta)$ *und* $\psi(q, \vartheta)$ *ist, wenn nicht alle Ableitungen zweiter Ordnung verschwinden.*)

Ein ganz ähnliches Ergebnis gilt, falls $M < 1$, für Verzweigungssingularitäten. Doch gehen wir auf diese nicht ein.

4. Grenzlinien L_1 und L_2. In der Diskussion des Verschwindens von J ist es offenbar von Vorteil, den Ausdruck für J gemäß (16) in seine Bestandteile zu zerlegen. Dementsprechend betrachten wir zu einer gegebenen Strömung den Ort $h_1(\xi, \eta) = 0$ in der ξ, η-Ebene und seine Abbildung auf die x, y-Ebene. Wir nehmen an, daß α weder $0°$ noch $90°$ ist. Sei nun p_0 ein Punkt mit Koordinaten ξ_0, η_0, für den $h_1(\xi_0, \eta_0) = 0$, $\left(\frac{\partial h_1}{\partial \eta}\right)_{\xi_0, \eta_0} \neq 0$ ist. Dann gibt es eine Kurve $\eta = g(\xi)$, $\eta_0 = g(\xi_0)$ längs der $h_1(\xi, \eta) = 0$. Wir nennen sie l_1, und ihr Bild L_1 in der x, y-Ebene *Grenzlinie*. (Natürlich ist l_1 selbst Verzweigungslinie der ξ, η-Ebene. Doch beziehen wir uns in der Namengebung nur auf die x, y-Ebene.) — Ein Punkt P der x, y-Ebene, für dessen Bild p die Gleichung $h_1(\xi, \eta) = 0$ erfüllt ist, heißt *Grenzpunkt.*

Sei m ein Punkt von l_1 und dort sowohl $\frac{\partial h_1}{\partial \xi}$ wie $\frac{\partial h_1}{\partial \eta}$ ungleich Null. Daher ist nach (22) $h_2 \neq 0$. Also hat nach (15) L_1 in M, dem Bilde von m die η-Richtung. Dasselbe gilt für jede Kurve C durch M, deren Bild c durch m etwa $\eta = f(\xi)$, dort *nicht* die ξ-Richtung hat, die wir als *Ausnahmerichtung* bezeichnen. Alle solchen Kurven, insbesondere die C^+ durch M, berühren die L_1 in M. Das gilt auch im Besonderen von den *Kurven konstanter Geschwindigkeit* q *und den Kurven konstanter Richtung* ϑ. Denn z.B. für letztere ist $d\vartheta = 0$, aber wegen $\alpha \neq 90°$ und (16): $d\eta = \frac{1}{2}(dQ + d\vartheta) = \frac{1}{2} dQ \neq 0$. Daher

berühren alle diese Kurven die Kurve L_1 in M und haben dort in jedem Punkt die C^+-Richtung.

Hat aber c in m die Ausnahmerichtung, so ist $\frac{d\eta}{d\xi} = f'(\xi) = 0$. Man kann nicht mehr schließen, daß die entsprechende C in M die η-Richtung hat, da nun $d\xi$ und $d\eta$ von gleicher Ordnung klein sind. Man sieht dann leicht, daß in M sowohl $\frac{dx}{d\xi} = \frac{\partial x}{\partial \xi} + \frac{\partial x}{\partial \eta} f'(\xi)$ wie auch $\frac{dy}{d\xi}$ verschwinden, während $\frac{d^2x}{d\xi^2}$ und $\frac{d^2y}{d\xi^2}$ nicht beide verschwinden, und daß folglich C *in M eine Spitze hat*. Insbesondere hat daher die Kurve C^- in M eine Spitze. Ebenso haben *Stromlinien* und *Potentiallinien* jede eine Spitze in M. Denn nach (17) folgt aus $h_1 = 0$, daß $\psi_\xi = 0$ ist, und da $d\psi = \psi_\xi\, d\xi + \psi_\eta\, d\eta = \psi_\eta\, d\eta = 0$, und da $\psi_\eta \neq 0$, wegen $h_2 \neq 0$, ist $d\eta = 0$ in m. Daher hat das Bild der

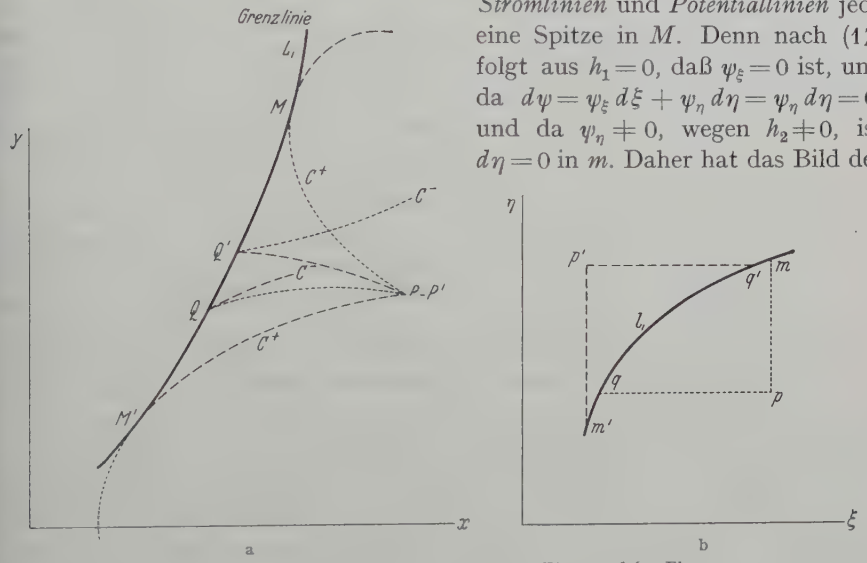

Fig. 1 a u. b. Umgebung der Grenzlinie in x, y-Ebene und ξ, η-Ebene

Stromlinie in m die Ausnahmerichtung und die Stromlinie in M eine Spitze. Dies ist in Übereinstimmung mit (20), wonach $\varkappa$ an der Grenzlinie unendlich wird. Nach (20) *wird an einer Grenzlinie auch die Beschleunigung unendlich.*

Wir fassen zusammen: *Wir betrachten in der ξ, η-Ebene den Ort der Punkte $h_1(\xi, \eta) = 0$, im allgemeinen eine Kurve l_1; wir betrachten Punkte m auf l_1, für die $\frac{\partial h_1}{\partial \xi} \neq 0$, $\frac{\partial h_1}{\partial \eta} \neq 0$ ist. Das Bild L_1 von l_1, die Grenzlinie, ist in allen solchen Punkten Einhüllende von C^+-Charakteristiken, von Kurven konstanter Richtung und Kurven konstanter Geschwindigkeit. Die C^--Charakteristiken, die Stromlinien und die Potentiallinien haben Spitzen auf L_1.* Für $h_2 = 0$ und die entsprechende Grenzlinie L_2 sind die Rollen von ξ und η vertauscht.

An der Grenzlinie stoßen zwei Lösungen zusammen (Fig. 1). Betrachten wir die Korrespondenz der beiden Ebenen x, y und ξ, η in der Umgebung dieser Linie. Ein Punkt P in der x, y-Ebene ist das Bild zweier Punkte p und p'. Durch P gehen zwei C^+-Charakteristiken, die die L_1 berühren, und zwar in M und M', mit entsprechenden Punkten m und m' auf l_1. Auf der

η-Linie durch m liegt p, auf der durch m' liegt p'. Diese beiden Punkte liegen auf entgegengesetzten Seiten der Kurve l_1, da η entlang PM und PM' in entgegengesetztem Sinne variiert. Durch P gehen auch zwei C^--Charakteristiken PQ und PQ', die mit den Charakteristiken PM und PM' die Winkel $2\alpha_1$ und $2\alpha_2$ einschließen. Der Punkt q liegt auf der ξ-Linie durch p und auf l_1, und q' liegt auf der ξ-Linie durch p' und auf l_1.

Wir erhalten in üblicher Weise eine eineindeutige Abbildung, wenn wir jeder der beiden Seiten der l_1 ein Blatt der physikalischen Ebene zuordnen (in der Fig. 1 mittels verschiedener Stricharten angedeutet). Jedes solche Blatt trägt dann nur eine Strömung.

5. Spitzen der Grenzlinie[2]). Wir haben bisher angenommen, daß im betrachteten Punkte $\frac{\partial h_1}{\partial \xi} \neq 0$, $\frac{\partial h_1}{\partial \eta} \neq 0$ ist. Sei nun $h_1 = 0$, $\frac{\partial h_1}{\partial \eta} = 0$. Dann folgt aus (22), daß auch $h_2 = 0$, da für $\sin 2\alpha \neq 0$ auch $\frac{\alpha'}{Q'} - 1 \neq 0$. Also ist in einem solchen Punkt $h_1 = h_2 = 0$. Wir nennen einen solchen Punkt *zweifachen Grenzpunkt*. In einem Punkte mit $h_1 = h_2 = 0$ verschwinden alle vier Ableitungen $\frac{\partial x}{\partial \xi}$, usw. Er ist ein isolierter singulärer Punkt.

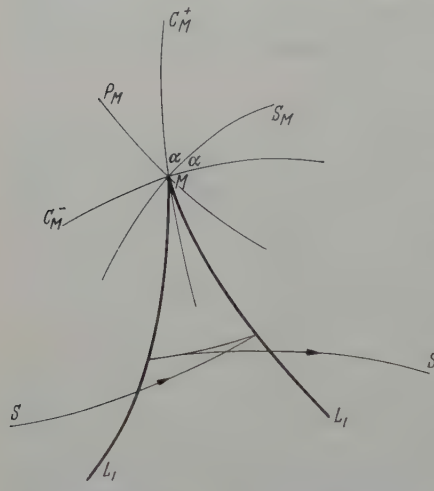

Fig. 2. Grenzlinie mit Spitze in M mit Charakteristiken C_M^+, C_M^-, Stromlinie S_M und Potentiallinie P_M in M. Stromlinie S mit Spitzen auf L_1

In m sei nun $h_1 = 0$, $\frac{\partial h_1}{\partial \eta} \neq 0$, $\frac{\partial h_1}{\partial \xi} = 0$, $\frac{\partial^2 h_1}{\partial \xi^2} \neq 0$. Dann hat die l_1-Kurve in m die ξ-Richtung. Schreiben wir $h_1(\xi, \eta) = 0$ als $\eta = g(\xi)$, so ist $g' = \frac{d\eta}{d\xi} = 0$ in m, aber $g'' \neq 0$ in m. Man zeigt dann, daß in m zwar $\frac{dx}{d\xi} = \frac{dy}{d\xi} = 0$ ist, daß aber $\frac{d^2x}{d\xi^2}$ und $\frac{d^2y}{d\xi^2}$ nicht beide verschwinden, *und daher L_1 in M eine Spitze hat.* Ferner sieht man, daß *die Kurve C^+ in M die Spitzentangente berührt, ohne selbst dort eine Spitze zu haben* (Fig. 2). Dasselbe gilt für jede Kurve durch M, deren Bild in m nicht die Ausnahmerichtung hat, also z. B. für die Kurven konstanter Richtung — damit meinen wir Kurven längs deren ϑ einen konstanten Wert hat — und die Kurven konstanter Geschwindigkeit. Die Kurve C^-, deren Bild die Ausnahmerichtung hat, schließt mit der Spitzentangente in M den Winkel 2α ein. Stromlinie und Potentiallinie in M halbieren diesen Winkel von innen und von außen. Die Bildkurve l_1 hat in m *ein Extremum*. Sie wird von der Stromlinie durch m in der ξ-Richtung *berührt*.

[2]) Dieser Teil fehlt ganz in den unter [4] und [7] zitierten Werken. In [3] und [5] sind die Resultate mit anderer Methode hergeleitet.

Diesen Spitzen der Grenzlinie, Extrema von l_1, kommt eine wichtige Rolle zu: Die Stromlinie, die durch die Spitze von L_1 geht, trennt diejenigen Stromlinien, die die Kurve L_1 nicht schneiden und völlig glatt verlaufen, von denen, die L_1 schneiden und daher an den Ästen von L_1 Spitzen haben (Fig. 2). Das Gebiet zwischen den Ästen von L_1 ist dreifach bedeckt. All dies kann z.B. an der bekannten RINGLEB-Strömung [1] verfolgt werden. Auch bei Fragen, die in Verbindung mit der heute so viel diskutierten „gemischten" Strömung um ein Profil auftreten, spielen die Spitzen der Grenzlinien eine Rolle.

6. Der Fall $M=1$[3]. Bis jetzt wurde $\alpha \neq 90°$, $M \neq 1$ angenommen. Wir betrachten nun den Fall $M=1$. Die Determinanten J und D verschwinden nun nicht mehr gleichzeitig. Aus (18) folgt, daß $D=0$, wenn $J<\infty$, und daß D jeden Wert annehmen kann, wenn $J\to\infty$. Da wir an der Hodographabbildung interessiert sind, wird D für uns maßgebend sein, und es ist von den in (18) gegebenen Formeln jetzt $D=\frac{\varrho}{q}\varphi_\xi\varphi_\eta$ zu verwenden, da in $D=\varrho q h_1 h_2 \cos^2\alpha$ für endliches D jetzt h_1 oder h_2 oder beide unendlich werden. Aus (17) folgt

$$\varphi_\xi=\frac{1}{\varrho}\left[-q\psi_q+\psi_\vartheta \operatorname{ctg}\alpha\right], \qquad \varphi_\eta=\frac{1}{\varrho}\left[q\psi_q+\psi_\vartheta\operatorname{ctg}\alpha\right]. \tag{23}$$

Da wir Verzweigungspunkte hier ausschließen, *setzen wir* $\varphi_\xi<\infty$, $\varphi_\eta<\infty$ *voraus*. Wir nennen die Linie $M=1$ *Schallinie* und bezeichnen sie mit S.

Ein Punkt von S heißt *gewöhnlich*, wenn dort $\varphi_\xi\neq 0$, $\varphi_\eta\neq 0$ ist. Es ist dann a fortiori $\psi_\xi\neq 0$, $\psi_\eta\neq 0$. Ein Beispiel einer *ganz aus gewöhnlichen Punkten bestehenden Schallinie* S ist z. B. der Kreis $M=1$ der x, y-Ebene bei der bekannten *Spiralströmung* oder Wirbelquelle[4]), die eine Superposition von Quelle und Wirbel darstellt. Sind C und K positive Konstante, so findet man

$$\left\{\begin{array}{ll} \varphi_\xi=K\cdot\frac{1}{\varrho}\operatorname{ctg}\alpha-C, & \varphi_\eta=K\cdot\frac{1}{\varrho}\operatorname{ctg}\alpha+C \\ \psi_\xi=-K+C\varrho\operatorname{tang}\alpha, & \psi_\eta=K+C\varrho\operatorname{tang}\alpha \\ \psi_q=C\frac{\varrho}{q}, \quad \psi_\vartheta=K, & b=\frac{K\varrho q^3}{K^2\operatorname{ctg}^2\alpha-C^2\varrho^2}. \end{array}\right. \tag{24}$$

Es ist also für $M=1$: $\psi_\vartheta\neq 0$, $\psi_q\neq 0$, $\varphi_\xi=-C$, $\varphi_\eta=C$, also alle Punkte gewöhnlich. Die Stromlinien beider Strömungen schneiden den Kreis $M=1$ unter einem von einem Rechten verschiedenen Winkel; *die Charakteristiken berühren diesen Kreis S daher nicht.*

An dem *Schallpunkt einer* L_1: $h_1=0$, ist $\psi_\xi=0$, und daher a fortiori $\varphi_\xi=0$, und analoges gilt für die L_2. Es kann gezeigt werden, daß ein Kurvenbogen L_1 oder L_2, längs dem überall $M=1$ ist, nicht möglich ist.

---

[3]) Diese Überlegungen sind im wesentlichen neu. Herrn G. S. S. LUDFORD bin ich für wertvollen Rat zu Dank verpflichtet.

[4]) Die hier benutzten Beispiele mögen in jedem der üblichen Lehrbücher, z.B. in dem bekannten Buche von SAUER, oder in den unter [4] oder [5] zitierten nachgelesen werden.

Eine neue Art von Grenzpunkt, wir nennen ihn *Schallgrenzpunkt*, typisch für $\alpha = 90°$, ist durch $\varphi_\xi = 0$, $\varphi_\eta = 0$, $\psi_\xi \neq 0$, $\psi_\eta \neq 0$, daher $h_1 \neq 0$, $h_2 \neq 0$ charakterisiert. Aus den ersten zwei Bedingungen folgt $\psi_q = 0$, $\psi_\vartheta \operatorname{ctg} \alpha = 0$. Also entweder $\psi_q = 0$, $\psi_\vartheta = 0$, oder $\psi_q = 0$, $\psi_\vartheta \neq 0$, aber so, daß $\psi_\vartheta \operatorname{ctg} \alpha \to 0$ mit $\alpha \to 90°$. Im ersten Fall handelt es sich um einen isolierten Ausnahmepunkt [vgl. (10) und (15′)]. Wichtiger ist der zweite Fall, da dieser offenbar entlang einer ganzen Kurve zutreffen kann, die wir eine *Schallgrenzkurve* L_t nennen. Das einfachste Beispiel einer Kurve L_t ist die *Radialströmung* (Quelle). Aus (24) für $C = 0$ ergibt sich, daß *für den ganzen Kreis* $M = 1$

$$\varphi_\eta = 0, \quad \varphi_\xi = 0, \quad D = 0, \quad b \to \infty, \quad h_1 \neq 0, \quad h_2 \neq 0. \tag{25}$$

In diesem Beispiel gibt es keine andere Grenzlinie, da h_1 und h_2 nach (17) überall endlich sind.

Gemäß unserer Definition ist jede Schallgrenzlinie eine Äquipotentiallinie. Daher ist die Stromlinienrichtung überall normal auf L_t, und da $\alpha = 90°$, so folgt, daß *sowohl die* C^+ *wie die* C^- *die Kurve* L_t *berühren.*

Kehren wir zu unserem Beispiel zurück, so sehen wir, daß *die Linien konstanter Richtung* ϑ *nun die* L_t *nicht berühren* (im Gegensatz zu unseren früheren Resultaten für die L_1 und L_2). Diese Linien sind nämlich die Radien, die auf der L_t senkrecht stehen und *mit den Stromlinien zusammenfallen*, und zwar mit den Stromlinien *beider Strömungen*. Jeder Punkt des Schallkreises gilt daher als ausgeartete Spitze.

Die Linie konstanter Richtung wird mit der Stromlinie lokal immer zusammenfallen, wo $\psi_q = 0$, $\psi_\vartheta \neq 0$. Denn aus $d\psi = \psi_q\, dq + \psi_\vartheta\, d\vartheta = \psi_\vartheta\, d\vartheta = 0$ folgt $d\vartheta = 0$. Dies kann, wie wir gesehen haben, an einer L_t zutreffen: $\psi_q = 0$, $\psi_\vartheta \neq 0$. Es kann auch für den Schallpunkt einer L_1 zutreffen, denn für $\alpha = 90°$ kann $\psi_\xi = -\psi_\vartheta + q\,\psi_q \operatorname{tang} \alpha$ verschwinden, wenn $\psi_q = 0$, $\psi_\vartheta \neq 0$. Dieselbe Formel zeigt aber, daß dies für $\alpha \neq 90°$ nicht möglich ist. Und in der Tat haben wir in **4** und **5** gefunden, daß an einer Kurve L_1 die Stromlinie mit der Linie konstanter Richtung den Winkel α macht.

Die aus der Definition von φ und ψ, gemäß (9), folgende Beziehung

$$dx = \frac{1}{q}\left(\cos\vartheta\, d\varphi - \frac{1}{\varrho}\sin\vartheta\, d\psi\right), \qquad dy = \frac{1}{q}\left(\sin\vartheta\, d\varphi + \frac{1}{\varrho}\cos\vartheta\, d\psi\right)$$

ergibt an der L_t, wo $d\varphi = 0$, $\psi_q = 0$

$$dx = -\frac{1}{\varrho\, q}\psi_\vartheta \sin\vartheta\, d\vartheta, \qquad dy = \frac{1}{\varrho\, q}\psi_\vartheta \cos\vartheta\, d\vartheta$$

$$\frac{dy}{dx} = -\cot\vartheta, \qquad \text{wenn} \quad d\vartheta \neq 0.$$

Das heißt, daß jedes Hodographelement mit $d\vartheta \neq 0$ in ein Element der x, y-Ebene senkrecht zur Stromrichtung abgebildet wird. Dies gilt, wie schon früher gesehen, für die L_t selbst, die eine Kurve konstanter Geschwindigkeit q und eine Potentiallinie ist, und für die beiden Charakteristiken C^- und C^+, während Stromlinie und Isokline, $d\vartheta = 0$, senkrecht auf der L_t stehen, da sie

im Hodograph die radiale Ausnahmerichtung aufweisen. Eine kleine Rechnung zeigt, daß solche Ausnahmekurven an der L_t *Spitzen* haben. Die Charakteristiken C^- und C^+ liegen auf verschiedenen Seiten der Isokline (bilden also miteinander einen Winkel von 180°), da ihre Bilder im Hodograph durch die Radien getrennt werden. Wir fassen zusammen: *Die gemäß 1.d) definierte Schallgrenzkurve, die eine Equipotentiallinie und eine Linie konstanter Geschwindigkeit ist, ist Einhüllende beider Scharen von Charakteristiken, und Ort der Spitzen von Stromlinien und Linien konstanter Richtung, die auf ihr senkrecht stehen.*

Da an einem Punkt mit $\psi_\xi = 0$ auch φ_ξ verschwindet, erkennt man, daß die folgende Klassifikation vollständig ist:

1.	$\psi_\xi \neq 0,\ \psi_\eta \neq 0$	a) $\varphi_\xi \neq 0,\ \varphi_\eta \neq 0$
		b) $\varphi_\xi = 0,\ \varphi_\eta \neq 0$
		c) $\varphi_\xi \neq 0,\ \varphi_\eta = 0$
		d) $\varphi_\xi = 0,\ \varphi_\eta = 0$
2.	$\psi_\xi = 0,\ \psi_\eta \neq 0$	a) $\varphi_\xi = 0,\ \varphi_\eta \neq 0$
		b) $\varphi_\xi = 0,\ \varphi_\eta = 0$
3.	$\psi_\xi \neq 0,\ \psi_\eta = 0$	a) $\varphi_\xi \neq 0,\ \varphi_\eta = 0$
		b) $\varphi_\xi = 0,\ \varphi_\eta = 0$
4.	$\psi_\xi = 0,\ \psi_\eta = 0$	$\varphi_\xi = 0,\ \varphi_\eta = 0$.

Hier charakterisiert 1.a) den *gewöhnlichen Punkt*, 1.d) *den Schallgrenzpunkt*, 2. und 3. den *Schallpunkt einer* L_1 *oder einer* L_2, 4. den *zweifachen Grenzpunkt.*

Wie früher bemerkt, können die Fälle 2., 3., 4. nur in isolierten Punkten auftreten. Dasselbe kann von 1.b) und 1.c) leicht eingesehen werden, so daß in der Tat die Schallgrenzlinie als einzige Grenz*linie* für $M = 1$ erscheint.

Wir betrachten noch die *Beschleunigung*

$$b = \frac{1}{2} q^3 \operatorname{tang} \alpha \cdot \left(\frac{1}{\varphi_\xi} + \frac{1}{\varphi_\eta} \right). \tag{20'}$$

Es geht also b an einer L_1, *an einer* L_2, *und an einer* L_t *gegen unendlich.* An einem regulären Punkt von S kann man keine Aussage machen, da $\varphi_\xi + \varphi_\eta = 0$ sein kann.

Ein interessantes Beispiel für den *Schallpunkt einer* L_1 (und einer L_2), bietet die erwähnte Ringleb-Strömung. Wir haben in diesem Falle (k eine Konstante):

$$\begin{cases} \psi_\xi = -\dfrac{k}{q} (\sin \vartheta \operatorname{tang} \alpha + \cos \vartheta), & \psi_\eta = -\dfrac{k}{q} (\sin \vartheta \operatorname{tang} \alpha - \cos \vartheta) \\ \varphi_\xi = \dfrac{k}{\varrho q} (\sin \vartheta + \cos \vartheta \operatorname{ctg} \alpha), & \varphi_\eta = -\dfrac{k}{\varrho q} (\sin \vartheta - \cos \vartheta \operatorname{ctg} \alpha), \end{cases} \tag{26}$$

und die Hodographengleichung der Grenzlinie lautet

$$\cos \vartheta = \pm \frac{1}{M} = \pm \sin \alpha.$$

Die Grenzlinie besteht aus zwei Teilen, einem oberhalb der x-Achse, entsprechend dem Ast $\cos\vartheta - 1/M = 0$, oder $\vartheta = 90° - \alpha$ $(0 \leqq \vartheta \leqq 90°)$; der andere unterhalb der x-Achse, entsprechend $\cos\vartheta + 1/M = 0$, oder $\vartheta = 90° + \alpha$ $(90° \leqq \vartheta \leqq 180°)$. Auf dem oberen ist für $\alpha = 90°$, $\psi_\xi \neq 0$, $\psi_\eta = 0$, daher ist dies eine L_2, deren Schallpunkt die Hodographkoordinaten $q = a$, $\vartheta = 0$ besitzt. Entlang dem anderen Ast ist für $\alpha = 90°$, $\psi_\xi = 0$, $\psi_\eta \neq 0$, also ist dies eine L_1, deren Schallpunkt die Koordinaten $q = a$, $\vartheta = 180°$ besitzt. [Für jeden dieser *beiden* Schallpunkte gilt $\varphi_\xi = 0$, $\varphi_\eta = 0$, also liegen die Fälle 2.b) und 3.b) obiger Klassifikation vor.] Diesen beiden verschiedenen Hodographpunkten entspricht *derselbe* Punkt in der x, y-Ebene, nämlich der Schnittpunkt A der Grenzlinie mit der x-Achse. Dieser Punkt ist *nicht ein zweifacher Grenzpunkt*, da er zwei verschiedenen Hodographenpunkten entspricht. Für jeden dieser beiden Punkte folgt, wegen $\varphi_\xi = 0$, $\varphi_\eta = 0$, $\psi_q = 0$, $\psi_\vartheta \neq 0$, aus dem früheren, daß die Grenzlinie in A beide Charakteristiken berührt, und daß die Richtung konstanten ϑ mit der Stromlinienrichtung zusammenfällt, was auch beides der Fall ist.

Die beiden Äste L_1 und L_2 schneiden sich im Unendlichen. *Dort* ist für $\alpha = 0°$, $\vartheta = 90°$ sowohl $h_1 = 0$ wie $h_2 = 0$, also ein zweifacher Grenzpunkt.

Auf weitergehende Diskussion von Grenzpunkten oder Grenzlinien gehen wir hier nicht ein. Wir wollten in dieser kleinen Arbeit die *Methode der Untersuchung* darlegen und zur Herleitung einiger Hauptresultate verwenden. Aus Raummangel muß auch eine Diskussion der Verzweigungslinien unterbleiben.

Literatur

[1] RINGLEB, F.: Exakte Lösungen der Differentialgleichungen einer adiabatischen Gasströmung. Z. angew. Math. Mech. **20**, 185—198 (1940). — [2] TOLLMIEN, W.: Grenzlinien adiabatischer Potentialströmungen. Z. angew. Math. Mech. **21**, 140—152 (1941). — [3] CRAGGS, J. W.: The breakdown of the hodograph transformation for irrotational compressible fluid flow in two dimensions. Proc. Cambridge Phil. Soc. **44**, 360—379 (1948). — [4] COURANT, R., and K. O. FRIEDRICHS: Supersonic Flow and Shock Waves. New York: Interscience Publishers, Inc. 1948. — [5] Modern Development in Fluid Dynamics High Speed Flow (particularly Chapter VII.) Edited by L. Howarth. Oxford University Press 1953. — [6] MEYER, R. E.: Focussing effects in twodimensional supersonic flow. Phil. Trans. Roy. Soc. Lond. **242**, 153—171 (1949). — [7] MANDEL, J.: Equilibres par tranches planes des solides à la limite d'écoulement. Paris: Louis Jean 1942. — [8] MISES, R. v.: Three remarks on the theory of the ideal plastic body. H. Reissner Anniversary Volume, p. 415—429. Edwards Brothers. Ann Arbor, Mich. 1949. — [9] GEIRINGER, H.: Some recent results in the theory of an ideal plastic body. Advances in Applied Mechanics, Vol. III, p. 197—294. New York: Academic Press 1953.

Harvard University, Cambridge (USA)

(Eingegangen am 3. Oktober 1955)

Mordell, L. J.
Math. Zeitschr. Bd. 63, S. 525—528 (1956)

The Minimum of an Inhomogeneous Quadratic Polynomial in n Variables

Dedicated to the memory of I. Schur

By
L. J. Mordell

Let

$$f(x) = f(x_1, x_2, \ldots, x_n) \tag{1}$$

be a real function of n real variables $(x) = (x_1, x_2, \ldots, x_n)$, and let (x') be a given set of real numbers. A well known problem is to find a number k independent of (x') such that there exist real numbers (x) with $(x) \equiv (x')$ (mod 1) and

$$f(x) \leqq k. \tag{2}$$

An alternative form of the question is to consider integers (x) such that

$$f(x + x') \leqq k. \tag{3}$$

Put in this form, we see that $f(x)$ may be replaced by an equivalent form $f(x) = g(y)$ arising from the unimodular integral substitution

$$x_l = \sum_{m=1}^{n} \lambda_{lm} y_m \qquad (l = 1, 2, \ldots, n).$$

Thus if we define (y') by

$$x'_l = \sum_{m=1}^{n} \lambda_{lm} y'_m,$$

the problem of making $f(x + x') \leqq k$ is equivalent to making $g(y + y') \leqq k$.

Let now

$$f(x) = \sum_{r,s=1}^{n} a_{rs} x_r x_s + 2 \sum_{r=1}^{n} a_r x_r, \tag{4}$$

where the coefficients are real, and $a_{rs} = a_{sr}$, $a_{rr} > 0$ for $1 \leqq r, s \leqq n$, and the determinant

$$A = \|a_{rs}\| \neq 0.$$

In the particular case when $n = 2$, and $a_1 = a_2 = 0$ and $f(x)$ is positive definite, Dirichlet, in a paper[1]) written in 1848, has shown that

$$4(a_{11} a_{22} - a_{12}^2)\, k = a_{11} a_{22} (a_{11} - 2 a_{12} + a_{22}) \tag{5}$$

[1]) Werke Bd. II, S. 29—48.

when $f(x)$ is replaced by the reduced equivalent form for which

$$a_{12} \geqq 0, \quad a_{11} \geqq a_{12}, \quad a_{22} \geqq a_{12}, \tag{6}$$

and that this is the best possible result.

I am not aware that any similar results have been found for $n \geqq 3$. By means of an exceedingly simple idea, I can include Dirichlet's result in the

Theorem. *Let k be defined by the determinantal equation*

$$\begin{vmatrix} a_{11}, & a_{12}, & \ldots, & a_{1n}, & \frac{1}{2}a_{11}+a_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}, & a_{n2}, & \ldots, & a_{nn}, & \frac{1}{2}a_{nn}+a_n \\ -\frac{1}{2}a_{11}+a_1, & \ldots, & -\frac{1}{2}a_{nn}+a_n, & k \end{vmatrix} = 0 \tag{7}$$

Then if the variables (x) can be divided into two sets of variables $(x_\lambda, x_{\lambda'}, \ldots)$ and $(x_\mu, x_{\mu'}, \ldots)$ such that all the cross products $a_{\lambda\lambda'}$ $(\lambda \neq \lambda')$, and $a_{\mu\mu'}$ $(\mu \neq \mu')$, are $\leqq 0$, then numbers (x) with assigned residues (mod 1) *exist such that*

$$f(x) \leqq k. \tag{8}$$

There is of course no real loss of generality in taking $a_r = 0$, and then the result takes the simpler form now given.

Let

$$F(x) = \sum_{r,s=1}^{n} A_{rs} x_r x_s,$$

where A_{rs} denotes the corresponding cofactor of a_{rs} in A. Then

$$4A\,k = F(a_{11}, a_{22}, \ldots, a_{nn}), \tag{9}$$

since now (7) becomes

$$\begin{vmatrix} a_{11}, & a_{12}, & \ldots, & a_{1n}, & a_{11} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}, & & \ldots, & a_{nn}, & a_{nn} \\ a_{11}, & & \ldots, & a_{nn}, & -4k \end{vmatrix} = 0.$$

When $n = 2$, (9) is the best possible result if and only if the conditions (6) hold, and then $A \geqq 0$.

We show first that on replacing (x) by $(x+\xi)$ for appropriate real (ξ), the inequality (8) takes the form

$$g(x) = \sum_{r=1}^{n} a_{rr}(x_r^2 - x_r) + \sum_{r \neq s} a_{rs} x_r x_s \leqq 0. \tag{10}$$

The condition for this is, on noting the new coefficients of x_r,

$$\begin{aligned} &a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n + a_1 + \tfrac{1}{2}a_{11} = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{n1}\xi_1 + \qquad\quad \cdots + a_{nn}\xi_n + a_n + \tfrac{1}{2}a_{nn} = 0, \\ &\qquad\qquad f(\xi_1, \xi_2, \ldots, \xi_n) \qquad\qquad = k. \end{aligned}$$

Multiply the first n equations in order by $-\xi_1, -\xi_2, \ldots, -\xi_n$, and add to the last. Then

$$(a_1 - \tfrac{1}{2}a_{11})\,\xi_1 + \cdots + (a_n - \tfrac{1}{2}a_{nn})\,\xi_n - k = 0.$$

Eliminating $\xi_1, \xi_2, \ldots, \xi_n$ from these equations, we have the equation (7) for k.

We show now that with the assigned conditions on the signs of the a, (10) is solvable with any assigned residues (mod 1) for (x). Clearly, we may without loss of generality suppose that $0 \leqq x_r < 1$ for $r = 1, 2, \ldots, n$; and we note equality occurs in (10) when $x_r = 0$. We prove that (10) is satisfied by at least one of the sets $(x_r - \varepsilon)$ where $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ and $\varepsilon_r = 0$ or 1. If this is not so, then for these 2^n sets of values, $g(x) > 0$. Take in particular

$$g(x_1, x_2, \ldots, x_n) > 0, \qquad g(x_1 - 1, x_2, \ldots, x_n) > 0. \tag{11}$$

We now apply an averaging process for x_1. Multiply the two inequalities in (11) by $1 - x_1$, x_1 respectively and add. Note that

$$x^2(1-x) + x(x-1)^2 = x(1-x), \qquad x(1-x) + (x-1)\,x = 0.$$

Write $X = x - x^2$. Then

$$a_{11}X_1 - a_{22}X_2 - \cdots - a_{nn}X_n + \sum_{r,s=1}^{r\neq s} a_{rs}x_r x_s > 0. \tag{12}$$

So averaging for x_2,

$$-a_{11}X_1 + a_{22}X_2 - \cdots - a_{nn}X_n + \sum_{r,s=2}^{r\neq s} a_{rs}x_r x_s > 0. \tag{13}$$

Suppose $n = 2$, then by addition, we have the contradiction $0 > 0$.

We prove now that (10) gives the best possible result, i.e. that $g(x) \geqq 0$ for all integers (x) if and only if

$$a_{12} \geqq 0, \quad a_{11} \geqq a_{12}, \quad a_{22} \geqq a_{12}. \tag{14}$$

These conditions are obviously necessary from $g(1, 1) \geqq 0$, $g(1, -1) \geqq 0$, $g(-1, 1) \geqq 0$. Next, they are sufficient. For

$$\begin{aligned} a_{11}(x_1^2 - x_1) + 2a_{12}x_1x_2 + a_{22}(x_2^2 - x_2) &\geqq a_{12}(x_1^2 - x_1 + 2x_1x_2 + x_2^2 - x_2) \\ &\geqq a_{12}(x_1 + x_2)(x_1 + x_2 - 1) \\ &\geqq 0. \end{aligned}$$

and the result follows.

The conditions (14) mean that here $A \geqq 0$ and that $a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ is a reduced form when $a_{12} \leqq \frac{1}{2}a_{11}$, and that

$$a_{11}(x_1 - x_2)^2 + 2a_{12}(x_1 - x_2)x_2 + a_{22}x_2^2$$

is when $a_{12} > \frac{1}{2}a_{11}$. This gives Dirichlet's theorem.

Suppose now $n \geqq 3$. We may suppose that $a_{rs} \leqq 0$, $r \neq s$, for $r \leqq m$, $s \leqq m$, say, and also $a_{rs} \leqq 0$ for $r > m$, $s > m$, $r \neq s$.

On averaging (12) for the variable $x_2, x_3, \ldots, x_m$ successively, we obtain

$$\sum_{r \leq m} a_{rr}X_r - \sum_{r > m} a_{rr}X_r + \sum_{r,s \neq 1,2,\ldots m}^{r \neq s} a_{rs}x_rx_s > 0. \tag{15}$$

So on averaging for the variables $x_{m+1}, \ldots, x_n$, we have

$$\sum_{r > m} a_{rr}X_r - \sum_{r \leq m} a_{rr}X_r + \sum_{r,s \neq m+1,\ldots n}^{r \neq s} a_{rs}x_rx_s > 0. \tag{16}$$

On adding (15), (16), we have a contradiction since the X's cancel and the coefficients occurring satisfy $a_{rs} \leqq 0$.

We note that when $n = 3$, the result holds if $a_{12} \leqq 0$ (or if $a_{23} \leqq 0$ or if $a_{31} \leqq 0$), and that when $n = 4$, the result holds if $a_{12} \leqq 0$, $a_{34} \leqq 0$.

The result, however, is not best possible since (14) shows that a necessary condition is obviously $a_{rs} \geqq 0$, $r, s = 1, \ldots, n$.

St. John's College, Cambridge, England

(Eingegangen am 13. Juni 1955)

Brauer, A.
Math. Zeitschr. Bd. 63, S. 529—541 (1956)

On the Schnirelmann density of the sum of two sequences

In memory of my teacher and friend I. Schur

By
Alfred Brauer

In 1936, Schur published his paper "Über den Begriff der Dichte in der additiven Zahlentheorie" [18]. Under the influence of this paper and discussions with Schur I wrote my papers [2] and [3]. The result of [3] will be improved by the present paper.

Let A be a sequence of positive integers $a_1 < a_2 < \cdots$ and $A(n)$ the number of elements $a_\nu \leqq n$. The greatest lower bound of the quotients $A(n)/n$ is called the Schnirelmann density α of A [17].

The sum $A + B$ of the sequences A and B is defined as the sequence of all integers of the form

$$e_1 a + e_2 b \quad (a \in A,\ b \in B;\ e_1 = 0 \text{ or } 1,\ e_2 = 0 \text{ or } 1).$$

We define the "modified" sum $A \dotplus B$ as the sequence of all integers of the form

$$a + e_2 b \quad (a \in A,\ b \in B;\ e_2 = 0 \text{ or } 1). \tag{1}$$

In the following, we denote the density of $A + B$ by γ, that of $A \dotplus B$ by γ', and that of B by β. It was conjectured by Landau [9, 10] that

$$\gamma \geq \min(1, \alpha + \beta). \tag{2}$$

This conjecture was studied by Khintchine [7], Besicovitch [1], Schur [18], and myself [2]. It was proved by Mann [12] in 1942. For the history of the problem, see the papers of Rohrbach [16], Erdös and Niven [6], Ostmann [15], and Niven [14].

The density of $A + B$ where $1 > \alpha > 0$ and $\beta = 0$ is not necessarily greater than α. Khintchine [8] called a sequence B of density zero an essential component if for every sequence A of positive density $\alpha < 1$ the density of $A + B$ is greater than α. He showed that the sequence of the squares is an essential component. More exactly, Khintchine proved:

If B is the sequence of the squares, then

$$\gamma \geq \alpha\{1 + \tfrac{1}{2} \cdot 10^{-8}(1 - \alpha)^2\}.$$

This result was improved by Buchstab [4] and Morimoto [13].

An important generalization was obtained by ERDÖS [5]. A sequence of positive integers is called a basic sequence of order h if every positive integer is the sum of at most h elements of the sequence. ERDÖS showed that every basic sequence is an essential component. (The converse is false; see LINNIK [11].) More exactly, ERDÖS proved:

If A is of density α and if B is a basic sequence of order h, then

$$\gamma \geq \gamma' \geq \alpha\left(1 + \frac{1-\alpha}{2h}\right). \tag{3}$$

Assume that B contains 1 as element. For each m let $g(m)$ be the smallest number of elements of B whose sum equals m. Assume that the smallest upper bound λ of the quotients $\left\{\sum_{m=1}^{n} g(m)\right\}/n$ is finite. STÖHR [20] calls λ the mean order of B.

LANDAU [10] showed that in (3) the order h can be replaced by the mean order λ. Moreover, he proved that $\lambda \leq \frac{19}{6}$ for the sequence of the squares.

LANDAU's result was improved by me in 1938 [3] when I proved that

$$\gamma \geq \gamma' \geq \alpha\left(1 + \frac{1-\sqrt{\alpha}}{\lambda}\right) \tag{4}$$

and by S. SELBERG [19] in 1944 who proved that

$$\gamma \geq \gamma' \geq \alpha\left\{1 + \frac{3(1-\alpha)}{4\lambda}\right\}. \tag{5}$$

(5) is better than (4) for $\alpha > \frac{1}{9}$ while (4) is better for $\alpha < \frac{1}{9}$.

In this paper these results will be improved further for $\frac{1}{9} \leq \alpha \leq \frac{1}{2}$. *It will be shown that for* $\alpha \leq \frac{1}{2}$

$$\gamma \geq \gamma' \geq \alpha + \tfrac{1}{2}\left\{\lambda - \sqrt{\lambda^2 - 3(\alpha - \alpha^2)}\right\}. \tag{6}$$

It is easy to see that (6) is better than (5) for $\frac{1}{9} \leq \alpha \leq \frac{1}{2}$ and better than (4) for $\alpha = \frac{1}{9}$. The case $\alpha > \frac{1}{2}$ will be studied in another paper.

The following example shows that (6) is sometimes better than (2). Let A be the sequence of the numbers which are congruent to 1 (mod 3), B the sequence consisting of the squares and of the numbers which are congruent to 2 (mod 90). Then the density of B equals $\frac{1}{90}$ and its mean order is less than or equal to $\frac{19}{6}$. It follows from (2) that $\gamma - \frac{1}{3} \geq \frac{1}{90}$ and from (6) that

$$\gamma - \frac{1}{3} \geq \frac{1}{2}\left(\frac{19}{6} - \sqrt{\frac{361}{36} - 1 + \frac{1}{3}}\right) = \frac{19 - \sqrt{337}}{12} > \frac{1}{20}.$$

In the following, we denote the sequence of the integers which are not elements of A by $\bar{A}$. For abbreviation, we set $A(n) = N$.

LEMMA 1 [SELBERG 19]. *We have*

$$\sum_{k=1}^{n} A(k)\{A(k) - A(k-1)\} = \tfrac{1}{2}(N^2 + N).$$

PROOF. We have

$$A(k)\{A(k)-A(k-1)\}=A(k) \quad \text{if} \quad k\in A,$$
$$A(k)\{A(k)-A(k-1)\}=0 \quad \text{if} \quad k\in \bar{A}.$$

Denoting the i-th element of A by a_i we obtain

$$\sum_{k=1}^{n} A(k)\{A(k)-A(k-1)\}=A(a_1)+A(a_2)+\cdots+A(a_N)$$
$$=1+2+\cdots+N=\tfrac{1}{2}(N^2+N).$$

LEMMA 2 [SELBERG 19]. *Let A be a sequence of positive integers with density $\alpha\leq\frac{1}{2}$. Assume*[1] *that n is an element of $\bar{A}$. We set $[n/2]=z$ and*

$$\bar{A}(n-z-1)=\varrho(1-\alpha)(n-z-1). \tag{7}$$

Let $D_m(n)$ be the number of elements $a_\varkappa$ of A for which $a_\varkappa\leq n-m$ and for which $a_\varkappa+m\in\bar{A}$. Then

$$2\sum_{m=1}^{z} D_m(n)\geq N-N^2+2\alpha N(n-z-1)+$$
$$+\alpha(2nz-z^2+\alpha z^2-\alpha n^2+2\alpha n-z-\alpha).$$

PROOF. An element k of A is counted in $\sum_{m=1}^{z} D_m(n)$ as often as there are numbers m such that $k+m\leq n$ and $k+m\in\bar{A}$. If $k\leq n-z-1$, then k must be counted $\{\bar{A}(k+z)-\bar{A}(k)\}$ times while an element k of A for which $n-z\leq k\leq n-1$ must be counted $\{\bar{A}(n)-\bar{A}(k)\}$ times. Hence

$$\begin{cases} \sum_{m=1}^{z} D_m(n)=\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}\{\bar{A}(k+z)-\bar{A}(k)\}+ \\ \qquad +\sum_{k=n-z}^{n-1}\{A(k)-A(k-1)\}\{\bar{A}(n)-\bar{A}(k)\}. \end{cases} \tag{8}$$

Since $\bar{A}(k)=k-A(k)$, we have

$$\bar{A}(k+z)-\bar{A}(k)=z-A(k+z)+A(k),$$
$$\bar{A}(n)-\bar{A}(k)=n-k-A(n)+A(k)=n-k-N+A(k),$$

hence by (8)

$$\begin{cases} \sum_{m=1}^{z} D_m(n)=\sum_{k=1}^{n-1}\{A(k)-A(k-1)\}A(k)+z\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}- \\ -\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)+(n-N)\sum_{k=n-z}^{n-1}\{A(k)-A(k-1)\}- \\ -\sum_{k=n-z}^{n-1} k\{A(k)-A(k-1)\}. \end{cases} \tag{9}$$

[1] This assumption is not necessary; it is made only to simplify the proof.

Since for any positive integers c and d with $c<d$

$$\sum_{k=c}^{d}\{A(k)-A(k-1)\}=A(d)-A(c-1),$$

it follows from (9) and Lemma 1 that

$$(10)\quad\begin{cases}\sum\limits_{m=1}^{z}D_m(n)=\frac{1}{2}(N^2+N)+zA(n-z-1)-\\ -\sum\limits_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)+(n-N)\{A(n-1)-A(n-z-1)\}-\\ -\sum\limits_{k=n-z}^{n-1}k\{A(k)-A(k-1)\}.\end{cases}$$

Now

$$\sum_{k=n-z}^{n-1}k\{A(k)-A(k-1)\}=-\sum_{k=n-z}^{n-1}A(k)+nA(n-1)-(n-z)A(n-z-1).$$

Substituting this in (10) and using $A(n-1)=A(n)=N$ we obtain

$$(11)\quad\begin{cases}\sum\limits_{m=1}^{z}D_m(n)=\frac{1}{2}(N-N^2)+NA(n-z-1)-\\ -\sum\limits_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)+\sum\limits_{k=n-z}^{n-1}A(k).\end{cases}$$

On the other hand

$$\begin{aligned}&\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\\ &\quad=\sum_{k=1}^{n-z-1}A(k+z)\{k-\bar{A}(k)-(k-1)+\bar{A}(k-1)\}\\ &\quad=\sum_{k=1}^{n-z-1}A(k+z)-\sum_{k=1}^{n-z-1}A(k+z)\{\bar{A}(k)-\bar{A}(k-1)\}\\ &\quad\leqq\sum_{k=1}^{n-z-1}A(k+z)-\alpha\sum_{k=1}^{n-z-1}(k+z)\{\bar{A}(k)-\bar{A}(k-1)\},\end{aligned}$$

$$(12)\quad\begin{cases}\sum\limits_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\leqq\sum\limits_{k=z+1}^{n-1}A(k)-\\ -\alpha\sum\limits_{k=1}^{n-z-1}k\{\bar{A}(k)-\bar{A}(k-1)\}-\alpha z\bar{A}(n-z-1).\end{cases}$$

Since $A(k)\geqq\alpha k$, we have

$$\bar{A}(k)=k-A(k)\leqq(1-\alpha)k,$$
$$k\geqq\bar{A}(k)/(1-\alpha).$$

It follows from (12) that

$$\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\leq\sum_{k=z+1}^{n-1}A(k)-$$
$$-\frac{\alpha}{1-\alpha}\sum_{k-1}^{n-z-1}\bar{A}(k)\{\bar{A}(k)-\bar{A}(k-1)\}-\alpha z\bar{A}(n-z-1)$$

and from Lemma 1 applied to $\bar{A}$ and from (7) that

$$\sum_{k=1}^{n-z-1}\{A(k)-A(k-1)\}A(k+z)\leq\sum_{k=z+1}^{n-1}A(k)-$$
$$-\frac{\alpha}{2(1-\alpha)}\{\bar{A}^2(n-z-1)+\bar{A}(n-z-1)\}-\alpha z\bar{A}(n-z-1)$$
$$=\sum_{k=z+1}^{n-1}A(k)-\frac{\alpha}{2}\{\varrho^2(1-\alpha)(n-z-1)^2+\varrho(n-z-1)+$$
$$-2\varrho z(1-\alpha)(n-z-1)\}.$$

Substituting this in (11) we obtain

$$(13)\quad\begin{cases}2\sum_{m=1}^{z}D_m(n)\geq N-N^2+2N\{n-z-1-\varrho(1-\alpha)(n-z-1)\}+\\+2\sum_{k=n-z}^{n-1}A(k)-2\sum_{k=z+1}^{n-1}A(k)+\alpha\varrho(n-z-1)\times\\\times\{\varrho(1-\alpha)(n-z-1)+1+2(1-\alpha)z\}.\end{cases}$$

Since $z=[n/2]$, we have $n-1\leq 2z$, hence $n-z\leq z+1$, and by (13)

$$(14)\quad\begin{cases}2\sum_{m=1}^{z}D_m(n)\geq N-N^2+2N(n-z-1)\{1-\varrho(1-\alpha)\}+2\sum_{k=n-z}^{z}A(k)+\\+\alpha\varrho(n-z-1)\{\varrho(1-\alpha)(n-z-1)+1+2(1-\alpha)z\}.\end{cases}$$

Denoting the right-hand side of (14) by $G(\varrho)$ we have

$$(15)\quad\begin{cases}\frac{\partial G}{\partial\varrho}=-2N(n-z-1)(1-\alpha)+2\alpha(1-\alpha)(n-z-1)^2\varrho+\\+\alpha(n-z-1)+2\alpha(1-\alpha)z(n-z-1).\end{cases}$$

Since $\partial G/\partial\varrho$ is a linear function of ϱ and since $\partial^2 G/\partial\varrho^2$ is positive, it follows that $G(\varrho)$ has one minimum and no maximum. If we denote the value of ϱ for which $G(\varrho)$ takes its minimum by ϱ_0, then it follows from (15) that

$$(16)\qquad\varrho_0=\frac{2N(1-\alpha)-\alpha-2\alpha(1-\alpha)z}{2\alpha(1-\alpha)(n-z-1)}\geq\frac{2n(1-\alpha)-1-2(1-\alpha)z}{2(1-\alpha)(n-z-1)}.$$

Now $\alpha\leq\frac{1}{2}$, hence $1\leq 2(1-\alpha)$, and by (16)

$$(17)\qquad\varrho_0\geq\frac{2n-2-2z}{2(n-z-1)}=1.$$

On the other hand, it follows from (7) that we have to consider only the interval $0 \leqq \varrho \leqq 1$. By (17), the function $G(\varrho)$ takes its smallest value in the interval $0 \leqq \varrho \leqq 1$ for $\varrho = 1$. Therefore by (14) if we replace $A(k)$ by αk and add

$$\begin{aligned} 2\sum_{m=1}^{z} D_m(n) &\geqq N - N^2 + 2\alpha N(n-z-1) + \alpha n(2z-n+1) + \\ &\quad + \alpha(n-z-1)\{(1-\alpha)(n+z-1)+1\} \\ &= N - N^2 + 2\alpha N(n-z-1) + 2\alpha n z - \alpha n^2 + \alpha n + \\ &\quad + (\alpha - \alpha^2)(n^2 - z^2 - 2n + 1) + \alpha n - \alpha z - \alpha. \\ 2\sum_{m=1}^{z} D_m(n) &\geqq N - N^2 + 2\alpha N(n-z-1) + \\ &\quad + \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha). \end{aligned}$$

LEMMA 3 [LANDAU *10*]. *Let A and B be sequences of positive integers and S the sequence of the numbers*

$$a + e_2 b \quad (a \in A,\ b \in B;\ e_2 = 0 \text{ or } 1).$$

Assume that B contains the number 1 as element. Let $g(m)$ be the smallest number of elements of B whose sum equals m. Assume that the mean order λ of B is finite. Using the notations of Lemma 2 we have

$$\sum_{m=1}^{z} D_m(n) \leqq \lambda z\{S(n) - N\}.$$

PROOF. We write m as sum of $g(m)$ elements of B

$$m = b_1 + b_2 + \cdots + b_{g(m)}.$$

Assume that $a \in A$ while $a + m \subset \bar{A}$ and $a + m \leqq n$. We consider the numbers

$$(18) \qquad a,\ a + b_1,\ a + b_1 + b_2, \ldots,\ a + b_1 + b_2 + \cdots + b_{g(m)} = a + m.$$

The first of these numbers is an element of A while the last is an element of $\bar{A}$. Hence there exists a smallest number f such that

$$a + b_1 + b_2 + \cdots + b_{f-1} \in A$$

while

$$s = a + b_1 + b_2 + \cdots + b_{f-1} + b_f \in \bar{A}.$$

Since s is the sum of an element of A and the element b_f of B, it is an element of S. Hence to each element a which must be counted in $D_m(n)$ belongs a sequence (18), hence at least one element s of S which does not belong to A. The number of such elements $s \leqq n$ is $S(n) - N$. Each such s can appear in at most $g(m)$ sequences (18) since it can be the second element, the third element, ..., the $\{g(m)+1\}$ st element of the sequence. Hence

$$(19) \qquad D_m(n) \leqq g(m)\{S(n) - N\},$$

$$(20) \qquad \sum_{m=1}^{z} D_m(n) \leqq \lambda z\{S(n) - N\}.$$

We now want to improve Lemma 3.

LEMMA 4. *Assume that* $n \in \bar{S}$. *We set* $z = [n/2]$. *Using the notations of Lemma 3 we have*

$$2\sum_{m=1}^{z} D_m(n) \leq (2\lambda z + 1)\{S(n) - N\} - \{S(n) - N\}^2.$$

PROOF. Let s_1 be an element of S which does not belong to A. This element will appear in exactly $g(m)$ sequences (18) if all the numbers

$$(21) \quad s_1 - b_1, \quad s_1 - b_1 - b_2, \quad s_1 - b_1 - b_2 - b_3, \quad \ldots, \quad s_1 - b_1 - b_2 - \cdots - b_{g(m)}$$

are elements of A. For certain values of s_1 and for certain values of m all the numbers (21) can belong to A. But the sequence (21) can contain elements of $\bar{A}$ or non-positive numbers.

In particular, let s_2 be another element of S which does not belong to A, and assume that

$$(22) \qquad 0 < s_1 - s_2 \leq z.$$

If we choose $m = s_1 - s_2$ and construct for s_1 and m the sequence (21), then the last element is s_2 which belongs to $\bar{A}$. Hence the sequence (21) contains at most $g(m) - 1$ elements of A, and s_1 appears in at most $g(m) - 1$ sequences (18) for $m = s_1 - s_2$.

Similarly, if $s \leq z$ is an element of S which does not belong to A, then for $m = s, s+1, \ldots, z$ the last element of the sequence (21) is not positive. Hence for such an m and s the element s appears in at most $g(m) - 1$ sequences (18).

For $1 \leq m \leq z$ we denote the number of pairs of elements s_1 and s_2 of S less than or equal to n which do not belong to A and for which $s_1 - s_2 = m$ by $\varphi(m)$ and the number of elements of S which do not belong to A and which are less than or equal to m by $\psi(m)$. Then it follows that there exist at least $\varphi(m) + \psi(m)$ elements of S for each m which do not belong to A and for which the last element of the sequence (21) is not in A. Hence we obtain instead of (19)

$$\begin{aligned} D_m(n) &\leq \{g(m) - 1\}\{\varphi(m) + \psi(m)\} + g(m)\{S(n) - N - \varphi(m) - \psi(m)\} \\ &= g(m)\{S(n) - N\} - \varphi(m) - \psi(m), \end{aligned}$$

and (20) can be replaced by

$$(23) \qquad \sum_{m=1}^{z} D_m(n) \leq \lambda z\{S(n) - N\} - \sum_{m=1}^{z}\{\varphi(m) + \psi(m)\}.$$

Let P_1 be the total number of pairs of elements of S less than or equal to n which do not belong to A and which satisfy (22), and P_2 be the number of such pairs for which

$$(24) \qquad s_1 - s_2 > z \geq \tfrac{1}{2}(n-1).$$

Then

$$(25) \qquad P_1 + P_2 = \tfrac{1}{2}\{S(n) - N\}^2 - \tfrac{1}{2}\{S(n) - N\}.$$

Moreover

$$P_1 = \sum_{m=1}^{z} \varphi(m). \tag{26}$$

We maintain that

$$P_2 \leqq \sum_{m=1}^{z} \psi(m). \tag{27}$$

Since $n \in \overline{S}$, it follows that $s_1 \leqq n-1$, hence by (24)

$$s_2 + z + 1 \leqq s_1 \leqq n - 1. \tag{28}$$

Therefore for a given s_2, there exist at most $n - 1 - s_2 - z$ values s_1 satisfying (24). Moreover, by (28)

$$s_2 \leqq n - 2 - z \leqq 2z - 1 - z = z - 1.$$

It follows that

$$P_2 \leqq \sum_{s_2 \leqq z-1} (n - 1 - s_2 - z). \tag{29}$$

On the other hand, we have to count s_2 in $\sum_{m=1}^{z} \psi(m)$ for $m = s_2, s_2+1, \ldots, z$, hence $z - s_2 + 1$ times,

$$\sum_{m=1}^{z} \psi(m) = \sum_{s_2 \leqq z} (z - s_2 + 1) \geqq \sum_{s_2 \leqq z-1} (z - s_2 + 1). \tag{30}$$

But

$$n - 1 - s_2 - z \leqq 2z - s_2 - z < z - s_2 + 1,$$

hence

$$\sum_{s_2 \leqq z-1} (n - 1 - s_2 - z) \leqq \sum_{s_2 \leqq z-1} (z - s_2 + 1), \tag{31}$$

and (27) follows from (29), (31), and (30).

Now we obtain from (26), (27), and (25)

$$\sum_{m=1}^{z} \{\varphi(m) + \psi(m)\} \geqq \tfrac{1}{2}\{S(n) - N\}^2 - \tfrac{1}{2}\{S(n) - N\}$$

and from (23)

$$2\sum_{m=1}^{z} D_m(n) \leqq (2\lambda z + 1)\{S(n) - N\} - \{S(n) - N\}^2.$$

LEMMA 5. *Using the same notations we have*

$$S(n) \geqq N + \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 + N^2 - 2N\alpha(n - z - 1) - N - \\ - \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha)\}^{\frac{1}{2}}. \tag{32}$$

PROOF. It follows from Lemma 4 that

$$S(n) - N \geqq \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 - 2\sum_{m=1}^{z} D_m(n)\}^{\frac{1}{2}}.$$

Now (32) follows from Lemma 2.

LEMMA 6. *The function*

$$f(x) = x + u - (v - wx + x^2)^{\frac{1}{2}}$$

where $v > 0$, $w > 0$, *and* $w^2 < 4v$ *is monotonic increasing.*

PROOF. We have

$$f'(x) = 1 - (x - \tfrac{1}{2}w)(v - wx + x^2)^{-\frac{1}{2}}.$$

If $x \leq \frac{1}{2}w$, then $f'(x)$ is positive. Assume now that $x > \frac{1}{2}w$. In order to prove that $f'(x)$ is positive it is sufficient to show that

$$(x - \tfrac{1}{2}w) < (v - wx + x^2)^{\frac{1}{2}}.$$

But this is correct since $w^2 < 4v$.

LEMMA 7. The inequality (32) remains correct if N is replaced by αn.

PROOF. We want to prove that the right-hand side of (32) satisfies the assumptions of Lemma 6 for $x = N$, $u = \lambda z + \frac{1}{2}$,

$$v = (\lambda z + \tfrac{1}{2})^2 - \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha), \tag{33}$$

$$w = 2\alpha(n - z - 1) + 1. \tag{34}$$

We distinguish between two cases.

I. *n is even.* Here $z = \frac{1}{2}n$. Since $\lambda > 1$, we have by (33)

$$4v > n^2 + 2n + 1 - 4\alpha n^2 + \alpha n^2 - \alpha^2 n^2 + 4\alpha^2 n^2 - 8\alpha^2 n + 2\alpha n + 4\alpha^2,$$

$$4v > n^2 + 2n - 1 - 3\alpha n^2 + 3\alpha^2 n^2 - 8\alpha^2 n + 2\alpha n + 4\alpha^2. \tag{35}$$

Moreover, by (34)

$$w^2 = \{\alpha(n - 2) + 1\}^2 = \alpha^2 n^2 - 4\alpha^2 n + 4\alpha^2 + 2\alpha n - 4\alpha + 1,$$

hence by (35)

$$4v - w^2 > n^2 + 2n - 3\alpha n^2 + 2\alpha^2 n^2 - 4\alpha^2 n + 4\alpha. \tag{36}$$

If we denote the right-hand side of (36) by $h(\alpha)$, then $\frac{\partial h(\alpha)}{\partial \alpha} = 0$ if

$$-3n^2 + 4\alpha n^2 - 8\alpha n + 4 = 0,$$

hence for $\alpha = \alpha_0$ where

$$\alpha_0 = \frac{3n^2 - 4}{4n^2 - 8n} > \frac{1}{2}$$

because $n > 2$. Since $\frac{\partial^2 h(\alpha)}{\partial \alpha^2}$ is positive, it follows that $h(\alpha)$ takes its smallest value in the interval $0 \leq \alpha \leq \frac{1}{2}$ for $\alpha = \frac{1}{2}$, hence by (36)

$$4v - w^2 > n^2 + 2n - \tfrac{3}{2}n^2 + \tfrac{1}{2}n^2 - n + 2 > 0,$$

hence $4v > w^2$ and $v > 0$.

II. *n is odd.* Here $z=\frac{1}{2}(n-1)$. It follows from (33) and (34) that

$$4v > n^2 - 4\alpha n^2 + 4\alpha n + \alpha n^2 - 2\alpha n + \alpha - \alpha^2 n^2 + 2\alpha^2 n - \\ - \alpha^2 + 4\alpha^2 n^2 - 8\alpha^2 n + 2\alpha n - 2\alpha + 4\alpha^2$$

$$= n^2 - 3\alpha n^2 + 3\alpha^2 n^2 + 4\alpha n - 6\alpha^2 n - \alpha + 3\alpha^2,$$

$$w^2 = \{\alpha(n-1)+1\}^2 = \alpha^2 n^2 - 2\alpha^2 n + \alpha^2 + 2\alpha n - 2\alpha + 1,$$

$$4v - w^2 > n^2 - 3\alpha n^2 + 2\alpha^2 n^2 + 2\alpha n - 4\alpha^2 n + \alpha + 2\alpha^2 - 1. \tag{37}$$

Let α_0 be the value of α for which the right-hand side of (37) takes its minimum, then

$$-3n^2 + 4\alpha_0 n^2 + 2n - 8\alpha_0 n + 1 + 4\alpha_0 = 0,$$

$$\alpha_0 = \frac{3n^2-2n-1}{4n^2-8n+4} = \frac{3n+1}{4n-4} > \frac{1}{2}$$

since $n>1$. It follows as above that the right-hand side of (37) takes its smallest value in the interval $0 \leqq \alpha \leqq \frac{1}{2}$ for $\alpha=\frac{1}{2}$, hence by (37)

$$4v - w^2 > n^2 - \tfrac{3}{2}n^2 + \tfrac{1}{2}n^2 + n - n + \tfrac{1}{2} + \tfrac{1}{2} - 1 = 0,$$

hence $4v > w^2$.

Therefore in both cases, the right-hand side of (32) satisfies the assumptions of Lemma 6. It follows from this Lemma that (32) remains correct if N is replaced by αn. We obtain

$$S(n) \geqq \alpha n + \lambda z + \tfrac{1}{2} - \{(\lambda z + \tfrac{1}{2})^2 - \alpha n + \alpha^2 n^2 - 2\alpha^2 n(n-z-1) - \\ - \alpha(2nz - z^2 + \alpha z^2 - \alpha n^2 + 2\alpha n - z - \alpha)\}^{\frac{1}{2}},$$

$$S(n) \geqq \alpha n + \lambda z - \tfrac{1}{2} - \{(\lambda z - \tfrac{1}{2})^2 - \alpha z(1-\alpha)(2n-z) - \alpha(n-z) + \alpha^2\}^{\frac{1}{2}}. \tag{38}$$

LEMMA 8 [SELBERG *19*]. *If B is a sequence of density $\beta \leqq \frac{1}{2}$, then the mean order λ of B is greater than or equal to $\frac{3}{2}$.*

PROOF. For the $B(n)$ numbers m which are elements of B and which are less than or equal to n, we have $g(m)=1$. For the remaining $n-B(n)$ numbers m which are less than or equal to n we have $g(m) \geqq 2$. Hence for $n=1, 2, \ldots$ it follows that $\lambda n \geqq B(n) + 2\{n - B(n)\}$,

$$B(n) \geqq (2-\lambda)n \qquad (n = 1, 2, \ldots). \tag{39}$$

On the other hand, if (39) holds, then the density β of B is

$$\beta \geqq 2 - \lambda,$$

hence

$$\lambda \geqq 2 - \beta \geqq 2 - \tfrac{1}{2} = \tfrac{3}{2}.$$

LEMMA 9. *For $0 \leqq \alpha \leqq \frac{1}{2}$ we have*

$$\alpha + \tfrac{1}{2}\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} < 1.$$

PROOF. If the Lemma is not correct, then

$$\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)} \geqq 2 - 2\alpha, \tag{40}$$

$$\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}$$

$$\geqq 2(1-\alpha)\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\},$$

$$3\alpha(1-\alpha) \geqq 2(1-\alpha)\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}$$

$$3\alpha \geqq 2\{\lambda + \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}.$$

This is impossible since $\lambda > 1$ and $\alpha \leqq \frac{1}{2}$. Hence (40) is not true, and the Lemma is proved.

Using these lemmas we can prove

THEOREM. *If A is a sequence of positive integers of density $\alpha \leqq \frac{1}{2}$, and B a sequence of mean order $\lambda \geqq \frac{3}{2}$, then the density γ' of the sequence S of the integers of the form*

$$a + e_2 b \quad (a \in A, \quad b \in B; \quad e_2 = 0 \text{ or } 1)$$

satisfies the inequality

$$\gamma' \geqq \alpha + \tfrac{1}{2}\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}. \tag{41}$$

By Lemma 8, the condition $\lambda \geqq \frac{3}{2}$ is satisfied if B is of density 0 or of positive density $\beta \leqq \frac{1}{2}$.

PROOF. We want to prove that for $n = 1, 2, \ldots$

$$S(n) > \alpha n + \tfrac{1}{2} n\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}. \tag{42}$$

This is trivial for $n = 1$ and $n = 2$ since 1 is in A, hence $2 \in S$, and since the right-hand side of (42) is less than n by Lemma 9. We now assume that (42) is already proved for all numbers less than n where $n \geqq 3$. Hence

$$S(n-1) > \alpha(n-1) + \tfrac{1}{2}(n-1)\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}.$$

If $n \in S$, then

$$S(n) = S(n-1) + 1 > \alpha(n-1) + \tfrac{1}{2}(n-1)\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} + 1$$

$$= \alpha n + \tfrac{1}{2} n\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\} + 1 - \alpha - \tfrac{1}{2}\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\},$$

hence by Lemma 9

$$S(n) > \alpha n + \tfrac{1}{2} n\{\lambda - \sqrt{\lambda^2 - 3\alpha(1-\alpha)}\}.$$

Therefore it is sufficient to prove (42) for the case that $n \in \bar{S}$.

We set $z = [n/2]$. Then the assumptions of the Lemmas 4, 5, and 7 hold. We have to distinguish between two cases.

I. *n is even.* It follows from (38) that

$$S(n) \geqq \alpha n + \tfrac{1}{2}(\lambda n + 1) - \tfrac{1}{2}\{(\lambda n + 1)^2 - 3\alpha(1-\alpha) n^2 - 2\alpha n + 4\alpha^2\}^{\frac{1}{2}},$$

$$S(n) \geqq \alpha n + \tfrac{1}{2}(\lambda n + 1) - \tfrac{1}{2}\{n^2(\lambda^2 - 3\alpha - 3\alpha^2) + 2n(\lambda - \alpha) + 1 + 4\alpha^2\}^{\frac{1}{2}}. \tag{43}$$

Now (42) follows from (43) if we can prove that

$$1-\{n^2(\lambda^2-3\alpha+3\alpha^2)+2n(\lambda-\alpha)+1+4\alpha^2\}^{\frac{1}{2}}>-\{n^2(\lambda^2-3\alpha+3\alpha^2)\}^{\frac{1}{2}}. \tag{44}$$

Hence we have to show that

$$\begin{aligned}
&1+n^2(\lambda^2-3\alpha+3\alpha^2)+2n(\lambda^2-3\alpha+3\alpha^2)^{\frac{1}{2}}\\
&\qquad>n^2(\lambda^2-3\alpha+3\alpha^2)+2n(\lambda-\alpha)+1+4\alpha^2,\\
&n(\lambda^2-3\alpha+3\alpha^2)^{\frac{1}{2}}>\lambda n-\alpha n+2\alpha^2,\\
&\lambda^2n^2-3\alpha n^2+3\alpha^2n^2>\lambda^2n^2+\alpha^2n^2+4\alpha^4-2\alpha\lambda n^2+4\alpha^2\lambda n-4\alpha^3 n,\\
&n^2(2\alpha\lambda-3\alpha+2\alpha^2)-4\alpha^2\lambda n+4\alpha^3 n-4\alpha^4>0.
\end{aligned}$$

Since $4\alpha^3 n-4\alpha^4>0$, it is sufficient to show that

$$n(2\lambda-3+2\alpha)-4\alpha\lambda\geqq 0.$$

The coefficient of n is greater than 0, hence it is sufficient that

$$n\geqq\frac{4\alpha\lambda}{2\lambda-3+2\alpha}=\frac{4\alpha}{2-(3-2\alpha)/\lambda}. \tag{45}$$

The right-hand side is a decreasing function of λ. Since $\lambda\geqq\frac{3}{2}$, it follows from (45) that $n\geqq 3$ is sufficient. But this is true. Therefore (44) is proved, and (42) holds if n is even.

II. *n is odd.* Here $z=\frac{1}{2}(n-1)$. It follows from (38) that

$$\begin{aligned}
S(n)&\geqq\alpha n+\tfrac{1}{2}(\lambda n-\lambda+1)-\tfrac{1}{2}\{(\lambda n-\lambda+1)^2-\\
&\quad-\alpha(n-1)(1-\alpha)(3n+1)-2\alpha(n+1)+4\alpha^2\}^{\frac{1}{2}},
\end{aligned}$$

$$\left\{\begin{aligned}
S(n)&\geqq\alpha n+\tfrac{1}{2}(\lambda n-\lambda+1)-\tfrac{1}{2}\{(\lambda n-\lambda+1)^2-3\alpha n^2+\\
&\quad+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2\}^{\frac{1}{2}}.
\end{aligned}\right. \tag{46}$$

We maintain that

$$\left\{\begin{aligned}
&-\lambda+1-\{(\lambda n-\lambda+1)^2-3\alpha n^2+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2\}^{\frac{1}{2}}\\
&\qquad>-(\lambda^2n^2-3\alpha n^2+3\alpha^2n^2)^{\frac{1}{2}}.
\end{aligned}\right. \tag{47}$$

Hence we have to prove that

$$\begin{aligned}
&-\lambda+1+\{\lambda^2n^2-3\alpha n^2+3\alpha^2n^2\}^{\frac{1}{2}}\\
&\qquad>\{(\lambda n-\lambda+1)^2-3\alpha n^2+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2\}^{\frac{1}{2}}.
\end{aligned}$$

The left-hand side is positive since

$$n^2(\lambda^2-3\alpha+3\alpha^2)\geqq n^2(\lambda^2-\tfrac{3}{4})>\lambda^2-1>(\lambda-1)^2.$$

Hence it is sufficient to prove that

$$\begin{aligned}
&\lambda^2-2\lambda+1+\lambda^2n^2-3\alpha n^2+3\alpha^2n^2-2(\lambda-1)(\lambda^2n^2-3\alpha n^2+3\alpha^2n^2)^{\frac{1}{2}}\\
&\quad>\lambda^2n^2+\lambda^2+1-2\lambda^2n+2\lambda n-2\lambda-3\alpha n^2+3\alpha^2n^2-2\alpha^2n-\alpha+3\alpha^2,\\
&\qquad 2\lambda^2n-2\lambda n+2\alpha^2n+\alpha-3\alpha^2>2(\lambda-1)n\{\lambda^2-3\alpha+3\alpha^2\}^{\frac{1}{2}}.
\end{aligned} \tag{48}$$

The left-hand side is greater than $2\lambda^2 n - 2\lambda n$ since $n \geqq 3$ while the right-hand side is less than $2\lambda^2 n - 2\lambda n$. Hence (48) and (47) are true, and (42) follows from (46) when n is odd.

Bibliography

[1] Besicovitch, A.: On the density of the sum of two sequences of integers. J. London Math. Soc. **10**, 246—248 (1935). — [2] Brauer, Alfred: Über die Dichte der Summe von Mengen positiver ganzer Zahlen. I. Ann. of Math. **39**, 322—340 (1938). — II. Ann. of Math. **42**, 959—988 (1941). — [3] Brauer, Alfred: Über die Dichte der Summe zweier Mengen, deren eine von positiver Dichte ist. Math. Z. **44**, 212—232 (1938). — [4] Buchstab, A.: Über ein metrisches Problem der additiven Zahlentheorie. Matematitscheskij Sbornik **40**, 190—195 (1933). — [5] Erdös, P.: On the arithmetical density of the sum of two sequences one of which forms a basis for the integers. Acta arithmetica **1**, 197—200 (1936). — [6] Erdös, P., and I. Niven: The $\alpha + \beta$ hypothesis and related problems. Amer. Math. Monthly **53**, 314—317 (1946). — [7] Khintchine, A.: Zur additiven Zahlentheorie. Matematitscheskij Sbornik **39**, 27—34 (1932). — [8] Khintchine, A.: Über ein metrisches Problem der additiven Zahlentheorie. Matematitscheskij Sbornik **40**, 180—189 (1933). — [9] Landau, E.: Die Goldbachsche Vermutung und der Schnirelmannsche Satz. Göttinger Nachrichten, Math.-phys. Kl. **1930**, 255—276. — [10] Landau, E.: Über einige neuere Fortschritte der additiven Zahlentheorie, S. 3—6 and 60—63. Cambridge 1937. — [11] Linnik, U. V.: On Erdös's theorem on the addition of numerical sequences. Matematitscheskij Sbornik, N. S. **10**, 67—78 (1942). — [12] Mann, H. B.: A proof of the fundamental theorem on the density of sums of sets of positive integers. Ann. of Math. (2) **43**, 523—527 (1942). — [13] Morimoto, S.: Über einen Satz von Khintchine für eine Summenfolge. Tôhoku Math. J. **43** 1—3, (1937). — [14] Niven, I.: The asymptotic density of integers. Bull. Amer. Math. Soc. **57**, 420—434 (1951). — [15] Ostmann, H.-H.: Über die Dichten additiv komponierter Zahlenmengen. Arch. Math. **1**, 393—401 (1848/49). [16] Rohrbach, H.: Einige neuere Untersuchungen über die Dichte in der additiven Zahlentheorie. Jber. dtsch. Math.-Ver. **48**, 199—236 (1938). — [17] Schnirelmann, L.: Über additive Eigenschaften von Zahlen. Math. Ann. **107**, 649—690 (1933). — [18] Schur, I.: Über den Begriff der Dichte in der additiven Zahlentheorie. Sitzgsber. preuß. Akad. Wiss. Phys.-math. Kl. **1936**, 269—297. — [19] Selberg, Sigmund: Note on a metrical problem in the additive theory of numbers. Arch. Math. Naturvidensk. **47**, 111—118 (1944). — [20] Stöhr, A.: Bemerkungen zur additiven Zahlentheorie. I. Mittlere Ordnung. J. reine u. angew. Math. **183**, 168—174 (1941).

University of North Carolina

(Eingegangen am 28. Oktober 1955)

Die Grundlehren der mathematischen Wissenschaften

in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete.

Herausgegeben von R. Grammel, E. Hopf, H. Hopf, F. Rellich, F.K. Schmidt, B. L. van der Waerden.

XXXIII. Band: **Algebra I.** Von Dr. **B. L. van der Waerden,** Professor der Mathematik an der Universität Zürich. Unter Benutzung von Vorlesungen von E. Artin und E. Noether. Vierte Auflage der Modernen Algebra, erster Teil. VIII, 292 Seiten Gr.-8°. 1955. Ganzleinen DM 29.60

Inhaltsübersicht: Zahlen und Mengen. — Gruppen. — Ringe und Körper. — Ganze rationale Funktionen. — Körpertheorie. — Fortsetzung der Gruppentheorie. — Die Theorie von Galois. — Unendliche Körpererweiterungen. — Reelle Körper. — Bewertete Körper. — Sachverzeichnis.

Aus den Besprechungen der dritten Auflage: Das Buch bietet eine glückliche Synthese der klassischen Algebra mit den modernen Begriffen der Gruppen-, Körper-, Bewertungs- und Idealtheorie sowie der Theorie der hyperkomplexen Zahlen. Die Darstellung ist klar und für die Einführung mit großem Geschick abgefaßt. Die drei ersten Kapitel dienen der Vorbereitung und enthalten die Grundbegriffe über Mengen, Gruppen, Ringe, Körper und Ideale. Dem schließt sich wesentlich die Theorie der kommutativen Körper an; sie wird bis zum Begriff der Bewertung eines Körpers geführt. Zahlreiche eingestreute Aufgaben unterstützen die Probe auf das Verständnis des Textes.

„Zeitschrift für angewandte Mathematik und Mechanik"

XXXIV. Band: **Algebra II.** Von Dr. **B. L. van der Waerden,** Professor der Mathematik an der Universität Zürich. Unter Benutzung von Vorlesungen von E. Artin und E. Noether. Dritte Auflage der Modernen Algebra, zweiter Teil. VIII, 224 Seiten Gr.-8°. 1955 Ganzleinen DM 29.60

Inhaltsübersicht: Eliminationstheorie. — Allgemeine Idealtheorie der kommutativen Ringe. — Theorie der Polynomideale. — Ganze algebraische Größen. — Lineare Algebra. — Theorie der hyperkomplexen Größen. — Darstellungstheorie der Gruppen und hyperkomplexen Systeme. — Sachverzeichnis.

Aus den Besprechungen der zweiten Auflage: ... Alles in allem ist das Studium des van der Waerden'schen Lehrbuches für den Kenner ein Genuß und für den Lernenden, man kann es wohl verantworten, „der Weg" in die moderne Algebra, der auch zugleich in die höchsten Höhen der Algebra führt. An sachlichen Kenntnissen wird nichts vorausgesetzt, nur eine straffe logische Denkweise ist erforderlich. Am Ende von jedem Kapitel kann der Lernende an den dort gestellten, sorgfältig ausgewählten Aufgaben prüfen, ob er den Text verstanden hat.

„Zeitschrift für mathematischen und naturwissenschaftlichen Unterricht"

SPRINGER-VERLAG / BERLIN · GÖTTINGEN · HEIDELBERG

Soeben erschien:

Randwertprobleme der Mikrowellenphysik

Von

Fritz E. Borgnis und **Charles H. Papas**

California Institute of Technology, Pasadena (California), USA

Mit 75 Textabbildungen

XV, 266 Seiten Gr.-8°. Ganzleinen DM 48.—

Inhaltsübersicht: 1. Das elektromagnetische Feld. Allgemeine und spezielle Beziehungen. — 2. Die skalare GREENsche Funktion. — 3. Beugung einer ebenen Welle am metallischen Kreiszylinder. Formulierung einer Integralgleichung für den Strombelag und Betrachtungen über deren Lösung. — 4. Formulierung einer stationären Darstellung des Streuquerschnitts für eine ebene Welle, deren elektrisches Feld parallel zur Achse des beugenden Kreiszylinders liegt. — 5. Der Streuquerschnitt bei der Beugung am metallischen Kreiszylinder für eine ebene Welle, deren elektrisches Feld senkrecht zur Zylinderachse liegt. — 6. Eine stationäre Darstellung für das Fernfeld bei der Beugung einer ebenen Welle am Kreiszylinder. — 7. Wahl einer Näherungsfunktion für den Strombelag am Kreiszylinder. — 8. Berechnung der Streuquerschnitte bei der Beugung am Kreiszylinder. — 9. Der Streuquerschnitt bei der Beugung einer ebenen Welle an einem unendlich langen metallischen Streifen und am unendlich langen Spalt. — 10. Die offen abstrahlende koaxiale Leitung mit Schirm. — 11. Unstetiger Übergang zwischen zwei koaxialen kreiszylindrischen Leitungen. — 12. Die kapazitive Blende im rechteckigen Hohlleiter. — 13. Die induktive Blende im rechteckigen Hohlleiter. — 14. Näherungsweise Bestimmung der Grenzfrequenz von zylindrischen Hohlleitern beliebigen Querschnitts. — 15. Näherungsweise Bestimmung der Eigenfrequenzen von Hohlraumresonatoren beliebiger Gestalt. — 16. Die Weitwinkel-Konusantenne. — 17. Die zylindrische Antenne. — 18. Anhang: A. Zum Zusammenhang zwischen der stationären Darstellung nach SCHWINGER und RAYLEIGHS Resultat für das Fernfeld bei der Beugung einer skalaren Welle an einer kleinen Scheibe. — B. Zur Beziehung zwischen Transmissionsfaktor und Fernfeld beim ebenen Schirm und deren Verwendung zur experimentellen Ermittlung des Transmissionsfaktors. — C. Grundzüge der Dyadenrechnung (Dyadenalgebra). — D. Die dyadische GREENsche Funktion. — Namen- und Sachverzeichnis.

SPRINGER-VERLAG / BERLIN · GÖTTINGEN · HEIDELBERG

Diesem Heft liegen zwei Prospekte des Springer-Verlages, Berlin - Göttingen - Heidelberg, bei.

Springer-Verlag, Berlin · Göttingen · Heidelberg. — Druck der Universitätsdruckerei H. Stürtz AG., Würzburg.
Printed in Germany.